

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 15, 2005, 15:49:28 ; Search time 40 Seconds  
(Without alignments)  
738.464 Million cell updates/sec

Title: US-09-896-791B-3  
Perfect score: 1636  
Sequence: 1 MAIGLQFVRVSNTELKEKSR.....TESSLPSWVLWALNPKCPCG 307  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_791,\*  
1: pir1,\*  
2: pir2,\*  
3: pir3,\*  
4: pir4,\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	802	49.0	667	JC7771	hypoxia inducible factor-3 alpha - human
2	503	30.7	811	JCT619	hypoxia-inducible
3	495	30.3	826	J38972	hypoxia-inducible
4	492	30.1	810	JCC4837	hypoxia-inducible
5	485	29.6	813	JCC5809	hypoxia-inducible
6	281	17.2	1505	JCC4851	hypoxia-inducible
7	256	15.6	248	A58520	hypoxia-inducible
8	202	12.3	655	A29945	single-minded gene
9	169.5	10.4	823	T21943	neurogenesis regul
10	169.5	10.4	825	T21944	hypothetical prote
11	142	8.7	805	JCT7635	hypothetical prote
12	138.5	8.5	392	JCT7633	aryl hydrocarbon r
13	138.5	8.5	791	A56241	aryl hydrocarbon r
14	136.5	8.3	776	A55448	Ah receptor nuclea
15	136.5	8.3	789	I59550	aryl hydrocarbon r
16	129.5	7.9	358	S58376	aryl hydrocarbon r
17	129.5	7.9	451	T42397	aryl hydrocarbon r
18	129	7.9	453	T19440	hypothetical prote
19	128	7.8	56	A41647	neurogenesis regul
20	125.5	7.7	626	JCC5405	brain and muscle A
21	123.5	7.5	626	JDE2072	Arnt-like PAS prot
22	118	7.2	701	JCC8065	aryl hydrocarbon r
23	117.5	7.2	834	JCT7993	aryl hydrocarbon r
24	105	6.5	848	S55514	aryl hydrocarbon r
25	104.5	6.4	1059	T30557	aryl hydrocarbon r
26	103.5	6.3	258	JCT619	brain and muscle A
27	103	6.3	846	JCT7721	aryl hydrocarbon r
28	103	6.3	1079	A70972	probable DNA polym
29	102.5			PC4288	brain and muscle A

juvenile hormone r  
aryl hydrocarbon r  
transcription fact  
hypothetical prote  
hypothetical nitra  
gene hindsight pro  
hypothetical prote  
aryl hydrocarbon r  
hemocytin - silk  
steroid hormone re  
aryl hydrocarbon r

Result No.	Score	Query	Match Length	DB ID	Description
1	49.0	667	2	JC7771	hypoxia inducible factor-3 alpha - human
2	68.0%	Best Local Similarity	68.0%	Pred. No. 4.9e-62;	Score 802; DB 2; Length 667;
3	172	Matches	172;	Conservative	Score 68.0%; Pred. No. 4.9e-62; Length 667;
4	11	Mismatches	28;	Indels	Score 68.0%; Pred. No. 4.9e-62; Length 667;
5	42	Gaps	5;		Score 68.0%; Pred. No. 4.9e-62; Length 667;
6	68	OY	9	RSTTELKEKRSRDAARSHRSQEETEVLYQLAHTLPFANGQVSAMHLDKASTMLTISYLRHR	#sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
7	66	Db	7	RSTTELKEKRSRDAARSHRSQEETEVLYQLAHTLPFANGQVSAMHLDKASTMLTISYLRHR	C;Species: Homo sapiens (man)
8	66	OY	69	LCAAGGGKGRATORILPFGPGGGFRHGTTRGRHLGLPVKGKQ-----QAPGQSV	C;Date: 01-Feb-2002
9	117	OY	69	LCAAGGGKGRATORILPFGPGGGFRHGTTRGRHLGLPVKGKQ-----QAPGQSV	C;Accession: JC7771
10	106	Db	67	LCAAGG-----EMVQGAGBEPDACYIKALEGFTVMVTAEGIMV	A;Title: Expression and characterization of hypoxia-inducible factor (HIF)-3alpha in huma
11	106	OY	118	DLCSSSLIINPTGCTNFS-LEIGHSTFDPIHRCQDQEELQDALTPRPLISKKKLEPTER	A;Reference number: JC7771; PMID:11573933
12	176	OY	118	DLCSSSLIINPTGCTNFS-LEIGHSTFDPIHRCQDQEELQDALTPRPLISKKKLEPTER	A;Contents: Kidney
13	176	Db	107	-----LERNVSKHGLGQSLELGHSTFDPIHRCQDQEELQDALTPQQLTSRKVEAPER	A;Molecule type: mRNA
14	160	Db	107	-----LERNVSKHGLGQSLELGHSTFDPIHRCQDQEELQDALTPQQLTSRKVEAPER	A;Residues: 1-667 <HAR>
15	220	OY	177	HFSLRLMKGSITLSRGRTLNKAATWKVLFICSGHMRAYKPAQTSPPASBRSERPLQCLVLI	C;Cross-references: UNIPROT:Q9Y2N7; DDBJ:AB054067
16	236	OY	177	HFSLRLMKGSITLSRGRTLNKAATWKVLFICSGHMRAYKPAQTSPPASBRSERPLQCLVLI	C;Comment: this protein is a heterodimeric transcription factor that belongs to the basic
17	220	Db	161	CFSLRMKGSITLSRGRTLNKAATWKVLFICSGHMRAYKPAQTSPPASBRSERPLQCLVLI	I've in the regulation of hypoxia-inducible gene expression in human kidney.
18	220	OY	237	CEAIPOLPHDGA 249	C;Genetics:
19	220	OY	237	CEAIPOLPHDGA 249	A;Gene: hif3-alpha
20	220	Db	221	CEAIP---HPGS 229	A;Map position: 19
21	220	OY	237	CEAIP---HPGS 229	C;Keywords: kidney

RESULT 2

JCT619

hypoxia-inducible factor 1 alpha - chicken

C;Species: Gallus gallus (chicken)

C;Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001





**QY** 184 STLTSRGRTNLKATWKVTLHCSHMR 210  
**C** Species: *Drosophila melanogaster* (fragment)  
**N**; Alternate names: single-minded gene protein  
**C**; Species: *Drosophila melanogaster*  
**C**; Accession: A23945  
**R**; Crews, S.T.; Thomas, J.B.; Goodman, C.S.  
**C**; Title: The *Drosophila* single-minded gene encodes a nuclear protein with sequence simil  
**A**; Reference number: A23945; MUID:88151023; PMID:3345560  
**A**; Accession: A23945  
**A**; Molecule type: mRNA  
**A**; Residues: 1-655 <CRE>  
**A**; Cross-references: UNIPROT:P05709; GB:M19020; NID:g158464; PID:g158465  
**C**; Genetics:  
**A**; Gene: sim  
**A**; Cross-references: FlyBase:FBgn004666  
**C**; Keywords: DNA binding; transcription regulation

Query Match 12.3%; Score 202; DB 2; Length 655;  
 Best Local Similarity 28.5%; Pred. No. 1.2e-09; Indels 70; Gaps 10;  
 Matches 73; Conservative 35; Mismatches 78; QY

36 QLAHTILPFGARGVS AHLDKASIMRITIS YLARMHRLLCAAGGKRGRATGRILLPEGGPFHGT 95  
 4 ELAKULPLPAAITSQLDKASIVLRLTYSKMRQYFPDG--LGRAWG---SSPAMORCAT 57

QY 96 HRR-GRHGL-----PGVKCQARGPQSYDVLCSSSLHNPTPGTNFSELEGHSI 143  
 58 IKGELSHIQTDLDRIFPVVAPPDKIMVYSETASTHGLS-----QVELTGNSI 105

QY 144 PDFHPCDQE B EAPL-----ERH 177  
 106 PEYIINYDQDEBNAILSHPHINOPLAQTHPTIGSPNQVQHPSAYDHRGSHTEIERT 165

QY 178 FSLMKTMSLTSGRTNLKATWKVTLHCSHMRAYKPPAQTSAGSPRSEPIQCLVLC 237  
 166 FFLRMKCVLAK--RNAGLRTSFGKVTHCGYKLARIYDPRGDGQS-----LIONGLELVA 218

QY 238 -----EAIPLPH 246  
 219 VGHSLPSSITEIKLUH 234

**Db**

**RESULT 9**

T21943 hypothetical protein F38A6.3a - *Caenorhabditis elegans*  
**C**; Species: *Caenorhabditis elegans*  
**C**; Accession: T21943; T26900  
**R**; Mortimore, B.  
**A**; Reference number: Z19491  
**A**; Accession: T21944  
**A**; Status: preliminary; translated from GB/EMBL/DDBJ  
**A**; Molecule type: DNA  
**A**; Residues: 1-825 <WIL>  
**A**; Cross-references: UNIPROT:Q9TMW0; EMBL:292833; PIDN:CA07381.1; GSPDB:GN00023; CESP:F314  
**A**; Experimental source: clone F38A6  
**R**; Rainscough, R.  
**A**; Submitted to the EMBL Data Library, June 1998  
**A**; Reference number: Z22082  
**A**; Accession: T26900  
**A**; Status: preliminary; translated from GB/EMBL/DDBJ  
**A**; Residues: 1-825 <WIL>  
**A**; Cross-references: EMBL:AU023842; PIDN:CAA19521.1; GSPDB:GN00023; CESP:F38A6.3b  
**C**; Genetics:  
**A**; Gene: CESP:F38A6.3b  
**A**; Experimental source: clone Y44A6D  

Query Match 10.4%; Score 16.9.5; DB 2; Length 825;  
 Best Local Similarity 26.3%; Pred. No. 1.1e-06; Indels 49; Gaps 10;  
 Matches 62; Conservative 35; Mismatches 90; QY

9 RSTELRKKEKS D A RRSR QSO TEV YQ L A M L PFA-RGV S H D A S M L T S Y R M H 67  
 110 KRMERERETSHAA DRRSR KES D IF D L K M C V P Y E G T W H D O R I A R T C R L 169

QY 68 RLCA--AGGKRGRATGRILLP R G P G F R H G T H R G R H G L P V G K C Q O A P G Q S V D C S S L 124  
 Do 170 KTGAGVLEA N D M E I T N E W T E D -----T R A E C D G - F V M V D S S S I L 212

QY 125 I R H P T P G N F S L -- E L G H S I P D F H P C D Q B E L Q A L T P R N L S K K L E A P -----T 174  
 Do 213 Y V T E S V A M V L G I T O D P U T G R A L R D F L H P S D Y D E F -----D K Q S K M H K P R G E D T D T 264

QY 175 E R H S L R M K S T L S R G T N L K A T W K -- V L H -- C S G H M R A Y K P A Q T S P A G 222  
 265 G I N N V L R M K T V I S P R G R C N L K S A L Y K S V S F L V H S K V S T G G H V S F M Q - G I T I P A G 318

**RESULT 10**

T21944 hypothetical protein F38A6.3b - *Caenorhabditis elegans*  
**C**; Species: *Caenorhabditis elegans*  
**C**; Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
**C**; Accession: T21944; T26900  
**R**; Mortimore, B.  
**A**; Submitted to the EMBL Data Library, March 1997  
**A**; Reference number: Z19491  
**A**; Accession: T21944  
**A**; Status: preliminary; translated from GB/EMBL/DDBJ  
**A**; Molecule type: DNA  
**A**; Residues: 1-825 <WIL>  
**A**; Cross-references: UNIPROT:Q9TMW0; EMBL:292833; PIDN:CA07381.1; GSPDB:GN00023; CESP:F314  
**A**; Experimental source: clone F38A6  
**R**; Rainscough, R.  
**A**; Submitted to the EMBL Data Library, June 1998  
**A**; Reference number: Z22082  
**A**; Accession: T26900  
**A**; Status: preliminary; translated from GB/EMBL/DDBJ  
**A**; Residues: 1-825 <WIL>  
**A**; Cross-references: EMBL:AU023842; PIDN:CAA19521.1; GSPDB:GN00023; CESP:F38A6.3b  
**C**; Genetics:  
**A**; Gene: CESP:F38A6.3b  
**A**; Map position: 5  
**A**; Introns: 19/2; 75/1; 11/2; 175/1; 226/3; 331/3; 518/1; 627/2; 713/1; 770/3  

Query Match 10.4%; Score 16.9.5; DB 2; Length 825;  
 Best Local Similarity 26.3%; Pred. No. 1.1e-06; Indels 49; Gaps 10;  
 Matches 62; Conservative 35; Mismatches 90; QY

9 R S T E L R K K E K S D A R R S R Q S O T E V Y Q L A M L P F A - R G V S H D A S M L T S Y R M H 67  
 110 K R M E R E T S H A A D R R S R K E S D I F D L K M C V P Y E G T W H D O R I A R T C R L 169

QY 68 RLCA--AGGKRGRATGRILLP R G P G F R H G T H R G R H G L P V G K C Q O A P G Q S V D C S S L 124  
 Do 170 KTGAGVLEA N D M E I T N E W T E D -----T R A E C D G - F V M V D S S S I L 212

QY 125 I R H P T P G N F S L -- E L G H S I P D F H P C D Q B E L Q A L T P R N L S K K L E A P -----T 174  
 Do 213 Y V T E S V A M V L G I T O D P U T G R A L R D F L H P S D Y D E F -----D K Q S K M H K P R G E D T D T 264

QY 175 E R H S L R M K S T L S R G T N L K A T W K -- V L H -- C S G H M R A Y K P A Q T S P A G 222

Db 265 GINMVILRMKTVISPRGRCLNKSALKYKVSVFLVSKVSVGHSFMQ-GITIPNG 318  
**RESULT 11**  
 N;Aryl hydrocarbon receptor nuclear translocator 1 - chicken  
 C;Species: Gallus gallus (chicken)  
 C;Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 09-Jul-2004  
 R;Citation: JC7635  
 A;Title: Hypoxia regulates avian cardiac Arnt and HIF-1alpha mRNA expression.  
 A;Reference number: JC7635; MUID:21294777; PMID:11401503  
 A;Accession: JC7635  
 A;Molecule type: mRNA  
 A;Residues: 1-805 <CAT>  
 A;Cross-references: UNIPROT:0985N3  
 C;Comment: This protein, known as hypoxia-inducible factor 1 beta, a member of the basic  
 genes, known as hypoxia-inducible factor 1 beta, a member of the basic  
 C;Genetics:  
 A;Gene: arnt1  
 C;Keywords: transcription factor  
 Query Match 8.7%; Score 142; DB 2; Length 805;  
 Best Local Similarity 20.5%; Pred. No. 0.00027;  
 Matches 61; Conservative 43; Mismatches 95; Indels 98; Gaps 8;  
 Db 14 LIRKEKSRRAAARSRSRQETEVLYQLAHTPPARGVSAHULKASIMRLTISYLHRHLCAAG 73  
 89 LARENHSIEERRRKMTAYITELSDMVPCTCSALARPKDPLKLTIRMASHM----- 139  
 Qy 74 GKRGRATGRPLPEGGGFRGHGRRGRHGLPVGKQCQAPGQPSUDLCSSLI---HNP 129  
 140 -KSLRGTTGNSTDGTYKPSLTDQBLKHLI---LEADGFLFIVSCENGRVVVSDFV 194  
 130 PGTNF-SLELIGHSTFDIHPCDQBLQDALTPRN-----LSKKKLEAPTERHF 178  
 195 PVLNQPOQSWFGSTLVDQHPDDPGKLRQLSNTSENALTEGTPKWCFLSNKDPAAPES-- 252  
 Db 179 SLRMKSTLTSRGRTNLKAATWK----- 201  
 253 -----ASKGRIDLKTCVKEKGQQSKRMCMSRSRSPICRMRCCGNSSDAVNRLS 304  
 Qy 202 -----VLACSGHMBAYKPAQTSAGSP-RSEEPQLCVLI 236  
 Db 305 FMRNRCRNGLGATKDGEPHYVWVHCTGYIKAWPPAGVSLPDDPDAGQSKFCLVAI 361  
**RESULT 12**  
 JCT633 aryl hydrocarbon nuclear translocator ARNT2X - zebra fish  
 C;Species: Brachydanio rerio (zebra fish)  
 C;Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 09-Jul-2004  
 C;Accession: JC7633  
 R;Hau, H.J.; Wang, W.D.; Hu, C.H.  
 Bloch, Biophys. Res. Commun., 282, 487-492, 2001  
 A;Title: Ectopic expression of negative ARNT2 factor disrupts fish development.  
 A;Reference number: JC7633; MUID:21294759; PMID:11401485  
 A;Content: Heart  
 A;Accession: JC7633  
 A;Molecule type: mRNA  
 A;Residues: 1-392 <HSU>  
 C;Comment: This factor, a basic helix-loop-helix PAS (bHLH-PAS) factor, heterodimerizes  
 C;Genetics:  
 A;Gene: arnt2X  
 Query Match 8.5%; Score 138.5; DB 2; Length 392;  
 Best Local Similarity 20.8%; Pred. No. 0.00023;  
 Matches 55; Conservative 48; Mismatches 103; Indels 59; Gaps 8;

Db 16 KEKRDAAARSRSRQETEVLYQLAHTPPARGVSAHULKASIMRLTISYLHRHLCAAGK 75  
**RESULT 13**  
 Qy 47 RENHSEIRRERRNKNTOTWELSDMVPCTCSALARPKDPLKLTIRMASHM-----K 96  
 Db 97 SMRGTTGNTSTDGAYKPSLTDQBLKHLI---LEADGFLFIVSCENGRVVSDSVFV 152  
 Qy 132 TNF-SLELIGHSTFDIHPCDQBLQDALTPRN-LSKKLEAPTE----- 170  
 Db 153 LNHRQSEWFGSTPEQVHDVDKLREQLSNTSENMSMTRGILDKTGTVKEGQSSMMRC 212  
 Qy 171 -----EAPTRHFSLARMKSTLTSRQRT---LNLKAAATKVLKGSCHGMAY 212  
 Db 213 MGSRSPFCRMRGCGSAPD-HTSLNRLASMRKYRNGLGPSSKEGEAQSVWHTGYIHW 271  
 Qy 213 KPPAQTSR-AGSPSPREPPQCLVIL 236  
 Db 272 PPAGMTIPDEDTEAQTSKYCLVAI 296  
**RESULT 14**  
 A55421 aryl hydrocarbon receptor nuclear translocator protein - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 09-Jul-2004  
 C;Accession: A55421  
 R;Reisz-Porszab, S.; Probst, M.R.; Fukunaga, B.N.; Hankinson, O.  
 Mol. Cell. Biol. 14, 6075-6086, 1994  
 A;Title: Identification of functional domains of the aryl hydrocarbon receptor nuclear translocator protein - mouse  
 A;Reference number: A55421; MUID:9434418; PMID:8065341  
 A;Accession: A55421  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-791 <REF>  
 A;Cross-references: UNIPROT:P53162; GB:U10325; NID:9555687; PID:AAA56717.1; PID:9555688  
 A;Note: authors translated the codon TTT for residue 375 as Ser, TCT for residue 380 as Ser  
 A;Note: authors failed to translate CAG for residue 507 as Gln  
 C;Genetics:  
 A;Gene: ARNT  
 Query Match 8.5%; Score 138.5; DB 2; Length 791;  
 Best Local Similarity 21.8%; Pred. No. 0.00053; Matches 58; Conservative 50; Mismatches 101; Indels 57; Gaps 9;  
 Db 14 LIRKEKSRRAAARSRSRQETEVLYQLAHTPPARGVSAHULKASIMRLTISYLHRHLCAAG 73  
 89 LARENHSIEERRRKMTAYITELSDMVPCTCSALARPKDPLKLTIRMASHM----- 139  
 Qy 174 GKRGRATGRPLPEGGGFRGHGRRGRHGLPVGKQCQAPGQPSUDLCSSLI---HNP 129  
 Db 140 -KSLRGTTGTGKPSLTDQBLKHLI---LEADGFLFIVSCENGRVVSDSVFV 194  
 Qy 130 PGTNF-SLELIGHSTFDIHPCDQBLQDALTPRN-LSKKLEAPTE----- 174  
 Db 195 PVLNQPOQSWFGSTLVDQHPDDPGKLRQLSNTSENALTEGTPKWCFLSNKDPAAPES-- 254  
 Qy 175 -----ERFLSMKSTL-----KGTNLKAAT----WKLHCSGMR 211  
 Db 255 MCMSRSRSPICRMRCCGTSVSDVPMNRFLNRCRNGLGSVKEGBPHFWVHCTGYIKA 314  
 Qy 212 YKPPAQTSR-AGSPSPREPPQCLVIL 236  
 Db 315 WPPAGMTIPDEDTEAQTSKYCLVAI 340

R.Li, H.; Dong, L.; Whitlock Jr., J.P.  
 J. Biol. Chem. 266, 28098-28105, 1994  
 A:Tittle: transcriptional activation function of the mouse Ah receptor nuclear translocat  
 A:Reference number: A55448; MUID:95050586; PMID:7961746  
 A:Accession: A55448  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residue: 1-776 <LIA>  
 A:Cross-references: UNIPROT:Q921F3; GB:U14333

Query Match 8.3%; Score 136.5; DB 2; Length 776;

Best Local Similarity 21.2%; Pred. No. 0.00077; Mismatches 109; Indels 57; Gaps 9;  
 Matches 58; Conservative 50; Mismatches 109; Indels 57; Gaps 9;

QY 6 QVRVSNTELKEKERAARSRSRSEOTEVYIQLAHTLPARGVSAHLDKASIMRLTISVL 65  
 Db 66 QMCNDKDERARENHEIERRRRNQNTAVITELSDMVPCSALARPKDQLTURMVAHM- 124  
 QY : ||| - ---HNPTPGTNF-SLELIGHSTDFIHDQEBELQDALITPREN-LSEKKEAPT--- 174  
 Db 66 MHRILCAAGGKRGRATGRGLPPEGCGPRHGCRHGLPVGKGQOAPERQSVICSSLI 125  
 125 - -----KSLRGTTGNTSTDGSYKPSFLIDQELKHLI---LIAADGFLFTIVSCETGRV 171  
 QY : ||| - ---ERHFSLRMKSTLS-----RGRTIANLKAT----WKVL 203  
 Db 232 EGQQSSMRMOMGSRSPICRMCCCTSSPVSMRSLFLRNRCNGLGSKGCEPHFVV 291  
 QY 204 HCSCSHMRAYKPPAOATSPACSPRS-EPPLOCLVLL 236  
 Db 292 HCTGYIKAWPPAGVSLPDDPEAQCGSKFCLVAI 325

## RESULT 15

139550 aryl hydrocarbon receptor nuclear translocator Arnt [imported] - human

C:Species: Homo sapiens (man)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004

R:Hoffman, E. C.; Reyes, H.; Chu, F.F.; Sander, P.; Conley, L.H.; Brooks, B.A.; Hankinson

Science 252, 954-958, 1991  
 A:Tittle: Cloning of a factor required for activity of the Ah (dioxin) receptor.

A:Reference number: 139550; MUID:91240280; PMID:1852076  
 A:Accession: 139550

A:Status: preliminary; translated from GB/EMBL/DDJB

A:Molecule type: mRNA  
 A:Residues: 1-789 <RES>

A:Cross-references: UNIPROT:P227540; GB:M69238; NID:9179003; PID:AAA51777.1; PID:9179004

A:Gene: GDB:ARNT  
 A:Cross-references: GDB:119701; OMIM:126110  
 A:Map position: 1q21-1q21

Query Match 8.3%; Score 136.5; DB 2; Length 789;  
 Best Local Similarity 21.8%; Pred. No. 0.00079; Mismatches 58; Conservative 50; Mismatches 101; Indels 57; Gaps 9;

QY 14 14 LIRKEKSRDAARRRSRQSOETVLYQLAHTLPARGVSAHLDKASIMRLTISYLHRLCAG 73  
 Db 89 LARENHSERRRKMTVYTTELMVPTCSALARPKDQLTILRMVASHM- 139  
 QY 74 KGRGATGTRGRPLGGFGFRHGRGRHGLPVGKCOQAPGPQSYVDLCCSSLI---HNPT 129  
 Db 140 -KSLRGTTGNTSTDGSYKPSFLIDQELKHLI---LEADGFLFTIVSCETGRV 194  
 QY 130 PGTFNP-SLRLIGHSIFDFIHPDCOBELODALTTPRN-LSKKKLRAPT----- 174  
 Db 195 PVLNQPOSEFWGSTIYDQVHPDDVKLREQLSSENALITGRLIDLKIGTVKKGQOSMR 254  
 QY 175 - ---ERHFSLRMKSTLS-----RPTLNKAAT-----WKVHSGMRA 211

Do 255 MONGSRESSIFICMRCCSSSVDPVSNRLSFVNRCRNLGLGSVTKDGEPFHVVWHTGYIKA 314  
 QY 212 YKPAQTSPPASPRS-BPPLQCVLI 236  
 Do 315 WPPAGVSLPDDPBPAGGSKFCLVAI 340

Search completed: June 15, 2005, 16:00:54  
 Job time : 42 secs

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

<b>OM protein - protein search, using sw model</b>										
Run on:	June 15, 2005, 15:37:58 ; Search time 173 Seconds (without alignments)									
Title:	US-09-896-791B-3									
Perfect score:	1636									
Sequence:	1 MAIGLQRVRSNTLRLKEKSR . . . . . TESSLPSWVILWALNRKNCPG 307									
<b>Scoring table:</b>	BLOSUM62									
Searched:	Gapop 10.0 , Gapext 0.5									
Total number of hits satisfying chosen parameters:	1612378									
Minimum DB seq length: 0										
Maximum DB seq length: 2000000000										
<b>Post-processing:</b>	Minimum Match 0% Maximum Match 100% Listing first 45 summaries									
<b>Database :</b>	UniProt_03;* 1: uniprot_sprot;* 2: uniprot_trembl;*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.										
<b>SUMMARIES</b>										
<b>Result</b>	<b>Score</b>	<b>Query</b>	<b>Match</b>	<b>Length</b>	<b>DB</b>	<b>ID</b>	<b>Description</b>			
No.										
1	1636	100.0	307	2	Q8VHR1		RESULT 1			
2	855.5	52.3	630	2	Q9QX54	Q8VHR1	PRELIMINARY;			
3	855.5	52.3	662	2	Q9Z115	Q8VHR1;	PRT; 307 AA.			
4	835	51.0	632	2	Q8WKA1	DT	01-MAR-2002 (TREMBlrel. 20, Created)			
5	835	51.0	659	2	Q6K72	DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)			
6	831	50.8	632	2	Q96K34	DT	01-Mar-2004 (TREMBlrel. 26, Last annotation update)			
7	820.5	50.2	662	2	Q9JHS2	DE	Inhibitory PAS domain protein.			
8	802	49.0	672	2	Q9Y2N7	GN	Name=Hif3a; Synonyms="tpas;			
9	518.5	35.0	648	2	Q9HA12	GS	Mus musculus (Mouse).			
10	518.5	31.7	811	1	HIF_A_CHICK	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Olfactory receptor.			
11	503	30.7	811	1	HIF_A_BOVIN	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; NCBI_TaxID=10090;			
12	499	30.5	823	1	HIF_A_BOVIN	ON	NCBI_TaxID=10090;			
13	499	30.5	823	2	Q8VHR1	RP	SEQUENCE FROM N.A.			
14	497	30.4	825	1	HIF_A RAT	RC	STRAIN=C57BL/6J; PubMed=11734056; DOI=10.1038/35107085;			
15	495	30.3	489	2	Q6EMI3	RA	Makino Y., Cao R., Svensson K., Bertilsson G., Asman M., Tanaka H., Cao Y., Berkenstam A., Poellinger L.;			
16	495	30.3	826	1	HIF_A_HUMAN	RT	"Inhibitory PAS domain protein is a negative regulator of hypoxia-inducible gene expression.";			
17	495	30.3	835	1	HIF_A_MOUSE	RL	Nature 414:550-551(2001).			
18	493	30.1	802	2	Q6PI54	RL	-1- SIMILARITY: Contains 1 basic helix-loop-helix (BHlH) domain.			
19	492.5	30.1	643	2	Q6STN6	CC	DR EML1; AR41641; HAL39015.1; -.			
20	492	30.1	819	2	Q7YES5	DR	MGD; MGI:1859778; Hif3a.			
21	490.5	30.0	777	2	Q6EHI4	DR	GO; GO_000534; C: nucleus; IC.			
22	489.5	29.9	774	2	Q6STN7	DR	GO; GO_000370; P:transcription factor activity; IPI.			
23	488	29.8	874	1	PAS1_MOUSE	DR	GO; GO_001666; P:response to hypoxia; IDA.			
24	488	29.8	874	2	Q6PEU2	DR	GO; GO_006566; P:transcription from Pol II promoter; IPI.			
25	484	29.6	821	2	Q64F54	DR	PFAM: PP00010; HIFH; 1.			
26	483.5	29.6	533	2	Q6P133	DR	SMART; SM0035; HIFH; 1.			
27	482	29.5	874	1	PAS1_RAT	DR	SEQUENCE 307 AA; 33990 MW; D89D3B225C9C3967 CRC64;			
28	481.5	29.4	626	2	Q6EGR9	DR	Q			
29	480	29.3	786	2	Q6SL11	DR	1 MAIGLQRVRSNTLRLKEKSR . . . . . TESSLPSWVILWALNRKNCPG 307			
30	479	29.3	835	2	Q69EW2	DR	1 MAIGLQRVRSNTLRLKEKSR . . . . . TESSLPSWVILWALNRKNCPG 307			
31	478	29.2	824	2	Q6H8T3	DR	1 MAIGLQRVRSNTLRLKEKSR . . . . . TESSLPSWVILWALNRKNCPG 307			
						QY	121 SSSLHNPPGTFNSLBLELIGHSHIFDPFDIPCDQEELQDLPALTPRNLSKKLEAPERHESL 180			
						QY	61 ISYLMRHLCAAGGGRKGRATGRLLPPEGFGFRGRHLGVPGKQOAPGPOSIDC 120			
						QY	121 SSSLHNPPGTFNSLBLELIGHSHIFDPFDIPCDQEELQDLPALTPRNLSKKLEAPERHESL 180			
						QY	61 ISYLMRHLCAAGGGRKGRATGRLLPPEGFGFRGRHLGVPGKQOAPGPOSIDC 120			
						QY	161 RMKSITLTSGRTNLKAATWKVUHCSGNRAVKPAQTSAGSPRSPLQCVLICAI 240			
						QY	181 RMKSITLTSGRTNLKAATWKVUHCSGNRAVKPAQTSAGSPRSPLQCVLICAI 240			
						QY	181 RMKSITLTSGRTNLKAATWKVUHCSGNRAVKPAQTSAGSPRSPLQCVLICAI 240			
						QY	241 POLPFHDGATLGLPOEKTPISTLFTPLWKLCLVLRKWPVQVQLQKGTESSLPSWVLM 300			

Db	241	PQLPFIDGAGIQRQKTPISLTFPDRMKWAKLICIVVNPVQVLOGKGTESSEPSWVIAL	300
Qy	301	NRKNCPG	307
Db	301	NRKNCPG	307
RESULT 2			
Q9QX54		PRELIMINARY;	PRT; 630 AA.
ID	Q9QX54;		
AC	Q9QX54;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)		
DE	Hypoxia-inducible factor 3 alpha (Fragment).		
GN	Name-Hif3a;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
NCBI_Taxid	10090;		
RN	[1]		
SEQUENCE FROM N.A.			
RA	Gu Y.-Z., Moran S.M., Hogenesch J.B., Wartman L., Bradfield C.A.; RT "Cloning and Characterization of a Third Hypoxia Inducible Factor, HIF3alpha.", J. Biol. Chem. 0:0-0(1999).		
RP	[2]		
SEQUENCE FROM N.A.			
RA	Gu Y.Z., Moran S.M., Hogenesch J.B., Wartman L., Bradfield C.A.; RT "Molecular characterization and chromosomal localization of a third alpha-class hypoxia inducible factor subunit, HIF3alpha.", Gene Expr. 7:205-213(1998).		
CC	-1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.		
DR	EMBL: AF079140; AAPF1782_1; JOINED.		
DR	EMBL: AF079141; AAPF1782_1; JOINED.		
DR	EMBL: AF079145; AAPF1782_1; JOINED.		
DR	EMBL: AF079147; AAPF1782_1; JOINED.		
DR	EMBL: AF079153; AAPF1782_1; JOINED.		
DR	EMBL: AF079140; AAPF1782_1; JOINED.		
DR	EMBL: AF079141; AAPF1782_1; JOINED.		
DR	EMBL: AF079143; AAPF1782_1; JOINED.		
DR	EMBL: AF079145; AAPF1782_1; JOINED.		
DR	EMBL: AF079147; AAPF1782_1; JOINED.		
DR	EMBL: AF079151; AAPF1782_1; JOINED.		
DR	EMBL: AF079152; AAPF1782_1; JOINED.		
DR	EMBL: AF079150; AAPF1782_1; JOINED.		
DR	EMBL: AF079148; AAPF1782_1; JOINED.		
DR	EMBL: AF079149; AAPF1782_1; JOINED.		
DR	EMBL: AF079151; AAPF1782_1; JOINED.		
DR	EMBL: AF079142; AAPF1782_1; JOINED.		
DR	HSSP; Q99814; IP97.		
MGD	MGI: 1839778; Hif3a.		
GO	GO:0005634; C:nucleus; IC.		
GO	GO:0001666; F:transcription factor activity; IPI.		
GO	GO:0001666; P:response to hypoxia; IDA.		
GO	GO:0006366; P:transcription from Pol II promoter; IPI.		
DR	EMBL: AF060194; AAC72734_1; -.		
DR	HSSP; Q99814; IP97.		
DR	MGD; MGI: 1839778; Hif3a.		
DR	GO; GO:0005634; C:nucleus; IC.		
DR	GO; GO:0001666; F:transcription factor activity; IPI.		
DR	GO; GO:0001666; P:response to hypoxia; IDA.		
DR	GO; GO:0006366; P:transcription from Pol II promoter; IPI.		
DR	InterPro; IPR001067; Nuc_translocat.		
DR	InterPro; IPR001610; PAC.		
DR	InterPro; IPR000014; PAS.		
DR	PFam; PF00010; HIF1.		
DR	PFam; PF00989; PAS; 1.		
DR	PRINTS; PR00785; NCTNSLOCATR.		
DR	InterPro; IPR001067; Nuc_translocat.		
DR	InterPro; IPR001610; PAC.		
DR	InterPro; IPR000014; PAS.		
DR	PFam; PF00010; HIF1.		
DR	PRINTS; PR00785; NCTNSLOCATR.		
DR	SMART; SM00353; HIF1.		
DR	SMART; SM00086; PAC; 1.		
DR	SMART; SM00091; PAS; 2.		
DR	PROSITE; PS50112; PAS; 2.		
DR	Pfam; PF00989; PAS; 1.		
DR	PRINTS; PR00785; NCTNSLOCATR.		
DR	SMART; SM00353; HIF1.		
DR	SMART; SM00086; PAC; 1.		
DR	SMART; SM00091; PAS; 2.		
FT	NON_TER 630 630		
SQ	SEQUENCE 630 AA; 69623 MW; 8288B2CB4B6D45B6 CRC64;		
Query Match	52.3%; Score 855.5; DB 2; Length 662;		
Best Local Similarity	68.4%; Pred. No. 2e-62;		
Matches	182; Conservative 5; Mismatches 24; Indels 55; Gaps 4;		
RESULT 3			
Q92215		PRELIMINARY;	PRT; 662 AA.
ID	Q92215		
AC	Q92215;		
DT	01-MAY-1999 (TREMBLrel. 10, Created)		
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)		
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)		
DE	Hypoxia inducible factor three alpha.		
GN	Name-Hif3a;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
NCBI_Taxid	10090;		
RN	[1]		
SEQUENCE FROM N.A.			
RA	Gu Y.Z., Moran S.M., Hogenesch J.B., Wartman L., Bradfield C.A.; RT "Molecular characterization and chromosomal localization of a third alpha-class hypoxia inducible factor subunit, HIF3alpha.", Gene Expr. 7:205-213(1998).		
RP			
SEQUENCE FROM N.A.			
RA	Gu Y.Z., Moran S.M., Hogenesch J.B., Wartman L., Bradfield C.A.; RT "Molecular characterization and chromosomal localization of a third alpha-class hypoxia inducible factor subunit, HIF3alpha.", Gene Expr. 7:205-213(1998).		
CC	-1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.		
DR	AF079153; AAPF1782_1; -.		
DR	AF079140; AAPF1782_1; JOINED.		
DR	AF079141; AAPF1782_1; JOINED.		
DR	AF079143; AAPF1782_1; JOINED.		
DR	AF079145; AAPF1782_1; JOINED.		
DR	AF079147; AAPF1782_1; JOINED.		
DR	AF079151; AAPF1782_1; JOINED.		
DR	AF079152; AAPF1782_1; JOINED.		
DR	AF079150; AAPF1782_1; JOINED.		
DR	AF079148; AAPF1782_1; JOINED.		
DR	AF079149; AAPF1782_1; JOINED.		
DR	AF079151; AAPF1782_1; JOINED.		
DR	AF079142; AAPF1782_1; JOINED.		
DR	HSSP; Q99814; IP97.		
DR	MGD; MGI: 1839778; Hif3a.		
DR	GO; GO:0005634; C:nucleus; IC.		
DR	GO; GO:0001666; F:transcription factor activity; IPI.		
DR	GO; GO:0001666; P:response to hypoxia; IDA.		
DR	GO; GO:0006366; P:transcription from Pol II promoter; IPI.		
DR	EMBL: AF060194; AAC72734_1; -.		
DR	HSSP; Q99814; IP97.		
DR	MGD; MGI: 1839778; Hif3a.		
DR	GO; GO:0005634; C:nucleus; IC.		
DR	GO; GO:0001666; F:transcription factor activity; IPI.		
DR	GO; GO:0001666; P:response to hypoxia; IDA.		
DR	GO; GO:0006366; P:transcription from Pol II promoter; IPI.		
DR	EMBL: AF060194; AAC72734_1; -.		
DR	InterPro; IPR001067; Nuc_translocat.		
DR	InterPro; IPR001610; PAC.		
DR	InterPro; IPR000014; PAS.		
DR	PFam; PF00010; HIF1.		
DR	PFam; PF00989; PAS; 1.		
DR	PRINTS; PR00785; NCTNSLOCATR.		
DR	SMART; SM00353; HIF1.		
DR	SMART; SM00086; PAC; 1.		
DR	SMART; SM00091; PAS; 2.		
DR	PROSITE; PS50112; PAS; 2.		
DR	Pfam; PF00989; PAS; 1.		
DR	PRINTS; PR00785; NCTNSLOCATR.		
DR	SMART; SM00353; HIF1.		
DR	SMART; SM00086; PAC; 1.		
DR	SMART; SM00091; PAS; 2.		
FT			
SQ	SEQUENCE 662 AA; 73012 MW; 58740A1B6993D3B5 CRC64;		
Query Match	52.3%; Score 855.5; DB 2; Length 662;		
Best Local Similarity	68.4%; Pred. No. 2e-62;		
Matches	182; Conservative 5; Mismatches 24; Indels 55; Gaps 4;		
QY	9 RSVTELKEKSRODARSRSQETEVLYQALHTLPARGYSAMHDKASINMLTISYLMR	68	
Db	7 RSVTELKEKSRODARSRSQETEVLYQALHTLPARGYSAMHDKASINMLTISYLMR	66	
Qy	69 LCAAGGKGGRATGRKLPEPGPGRHRGRHGLPVGKQCQPGPSVLDL-----	120	
Db	67 LCAAGGEMW-----QVEKGEPEPDACYLKALEGF	94	
Qy	7 RSVTELKEKSRODARSRSQETEVLYQALHTLPARGYSAMHDKASINMLTISYLMR	66	
Db	69 LCAAGGKGGRATGRKLPEPGPGRHRGRHGLPVGKQCQPGPSVLDL-----	120	
Qy	67 LCAAGGEMW-----QVEKGEPEPDACYLKALEGF	94	



Qy	229	PLOCVLVLCIAIPQLPFHGA	249	Query Match 50.8%; Score 811; DB 2; Length 632;
Db	215	PLOCVLVLCIAIP---HPGS	231	Best Local Similarity 68.2%; Pred. No. 2e-60; Indels 42; Gaps 5; Matches 178; Conservative 11; Mismatches 30;
RESULT 6				
ID	Q96K34	PRELIMINARY;	PRT;	632 AA.
AC	Q96K34;			
DT	01-DEC-2001	(TREMBlrel. 19, Created)		
DT	01-DEC-2001	(TREMBlrel. 19, Last sequence update)		
DE	O1-MAR-2004	(TREMBlrel. 26, Last annotation update)		
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TAXID=9606;				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Ovary; DOI=10.1038/ng1285; PubMed=14720339;			
RX	Ota T., Suzuki Y., Nisikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Ohayashi M., Nishi T., Tanaka T., Isini S., Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoi T., Kaku Y., Kocaira H., Kondo H., Sugawara M., RA			
RA	Takahashi M., Kanda K., Yokoji Y., Furuya T., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Niimura K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata M., Watanabe M., Hiroka S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Nisida M., Hotota T., Kubano J., RA			
RA	Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y., Toiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., RA			
RA	Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshikawa Y., Matsunaga H., Ichihara T., Shiohara N., Sano S., RA			
RA	Nakagawa S., Seroh A., Mizoguchi H., Gott Y., Shimizu P., Wakebe H., RA			
RA	Hishigaki H., Watanabe T., Sugiyama A., Takekoto M., Kawakami B., RA			
RA	Yamazaki M., Watanabe T., Kumagai A., Itakura S., RA			
RA	Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., RA			
RA	Kawabata A., Hikili T., Kobatake N., Inagaki H., Ikeda Y., Okamoto S., RA			
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., RA			
RA	Matsumura K., Nakajima Y., Mizuno T., Morinaga M., RA			
RA	Togashi T., Ogawa M., Hata H., Watanabe M., Komatsu T., RA			
RA	Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., RA			
RA	Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R., RA			
RA	Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S., RA			
RT	"Complete sequencing and characterization of 21,243 full-length human CC			
RT	-!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.			
RT	CDNA's.;"			
RT	Nat. Genet. 36:40-45 (2004).			
RL	"Complete sequencing and characterization of 21,243 full-length human DR			
CC	-!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.			
DR	EMBL; AL027725; BAB55324.1; -.			
DR	HSRP; Q9P814; 1P97.			
DR	GO; GO:0005634; C-nucleus; IEA.			
DR	GO; GO:000871; F-signal transducer activity; IEA.			
DR	GO; GO:0003700; P-transcription factor activity; IEA.			
DR	GO; GO:0006335; P-regulation of transcription, DNA-dependent; IEA.			
DR	GO; GO:0007165; P-signal transduction, IEA.			
DR	InterPro; IPR01052; HUH basic.			
DR	InterPro; IPR01067; Nuc_translocat.			
DR	InterPro; IPR00160; PAC.			
DR	InterPro; IPR00014; PAS.			
DR	Pfam; PF00010; HUH; 1.			
DR	Pfam; PF00999; PAS; 1.			
DR	PRINTS; PR00785; NCTRNSLOCATR.			
DR	SMART; SM0033; HUH; 1.			
DR	SMART; SM00066; PAC; 1.			
DR	SMART; SM00091; PAS; 2.			
DR	PROSITE; PS50112; PAS; 2.			
DR	SEQUENCE 662 AA; 72887 NM; AC9672E340544010 CRC64;			
PRINTS; PR00785; NCTRNSLOCATR.				
SMART; SM0033; HUH; 1.				
SMART; SM00066; PAC; 1.				
SMART; SM00091; PAS; 2.				
PROSITE; PS50112; PAS; 2.				
SEQUENCE 652 AA; 68933 MW; A19F1ED3D05F7A1 CRC64;				

7 RSTBLRKERSRDAARSRSQETEVLYOAHTRPARGVSADKASIMRLTISYLRER 66  
 69 LCAAGGKRGRATGRULPEPGPFGRHGRHGLPVGKCO-----QAPGPSV 117  
 67 LCAAG-----EWNOYRKESRPLACYLKALEGFVNVLTAEGDMAY 106  
 118 DLCCSSLINHPTPGNFS LEHIGHSIFDPHPCDQEELDALTPRPNUSKKLEAPTER 176  
 107 -----LSENVSKHGLSOLBLEIGHSFDPHPCDQEELDALTPRPSLSKGSEATER 160  
 177 HFSLRMKSTLTSRGETNLNIKAATKWLHSGHMRAYKPPAQTSPGSPRSEPPQCLVLI 236  
 161 HFSLRMKSTLTSRGETNLNIKAATKWLHSGHMRAYKPPAQTSPGSPRSEPPQCLVLI 220  
 237 CEAIP-----OLPFDGATL 251  
 221 CEAIPHSALPPLPGRGAFL 240

RESULT 8

Q9Y2N7 PRELIMINARY; PRT; 667 AA.  
 ID Q9Y2N7  
 AC 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Putative homolog of hypoxia inducible factor three alpha (Hypoxia-inducible factor-3 alpha).  
 DR Name=HIF-3A;  
 OS Homo sapiens (Human);  
 OC Buxaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RA Lamerdin J.B., McCreedy P.M., Skowronski E., Viswanathan V.,  
 RA Burkhardt-Schultz K.J., Gordon L., Dias J., Ramirez M., Stilwagen S.,  
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Barnes J., Liu S.,  
 RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J.,  
 RA Attix C., Andreise T., Tranklein M., Amico-Keller G., Coefield J.,  
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,  
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
 RA Olsen A.S., Carrasco A.V.,  
 RL Submitted (APR 1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=kidney;  
 RX MEDLINE=21458277; PubMed=11573933; DOI=10.1007/bfrc.2001.5659;  
 RA Hara S., Hamada J., Kobayashi C., Kondo Y., Imura N.;  
 RT "Expression and characterization of hypoxia-inducible factor (HIF)-1alpha in human kidney: suppression of HIF-mediated gene expression by HIF-3alpha.",  
 RT Bloemberg, Biophys. Res. Commun. 287:808-813(2001).  
 CC -- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
 DR EMBL: AC001193; ADD2668.1; -.  
 DR EMBL: AB054067; BAB69689.1; -.  
 PIR: JC7771; JC7771.  
 HSSP: Q1665; 1QOB.  
 GeneW: HGNC:15825; HIF3A.  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0005634; C:nuclear transducer activity; IEA.  
 DR GO: GO:0003700; P:transcription factor activity; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR0102; Hifc basic.  
 DR InterPro: IPR01067; Nuc\_translocat.  
 DR InterPro: IPR001610; PAC.  
 DR InterPro: IPR00014; PAS.  
 DR Pfam: PF00989; PAS; 1.  
 DR PRINTS; PR00785; NOTRNSLOCATR.  
 SMART; SM00086; PAC; 1.  
 SMART; SM00031; PAS; 2.  
 PROSITE; PS00112; PAS; 2.  
 SEQUENCE; 667 AA; 72404 MN; 67B8794FF9DCCF4B CRC64;

DR SMART; SM00086; PAC; 1.  
 DR SMART; SM00031; PAS; 2.  
 DR PROSITE; PS00112; PAS; 2.  
 DR SEQUENCE; 667 AA; 72404 MN; 67B8794FF9DCCF4B CRC64;

Query Match Best local similarity 68.0%; Pred. No. 5..5e-58; Score 802; DB 2; Length 657; Matches 172; Conservative 11; Mismatches 28; Indels 42; Gaps 5;

Q9Y2N7 9 RSTBLRKERSRDAARSRSQETEVLYOAHTRPARGVSADKASIMRLTISYLRER 68  
 177 HFSLRMKSTLTSRGETNLNIKAATKWLHSGHMRAYKPPAQTSPGSPRSEPPQCLVLI 236  
 161 HFSLRMKSTLTSRGETNLNIKAATKWLHSGHMRAYKPPAQTSPGSPRSEPPQCLVLI 220  
 237 CEAIP-----OLPFDGATL 251  
 221 CEAIPHSALPPLPGRGAFL 240

RESULT 9

Q9HAI2 PRELIMINARY; PRT; 648 AA.  
 ID Q9HAI2  
 AC 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE Hypothetical Protein FIJ1191.  
 OS Homo sapiens (Human).  
 OC Buxaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC TISSUE=whole embryo;  
 RX PubMed=14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahori K.,  
 RA Murakami K., Yasuda T., Iwayanagi T., Watanabe M., Shiratori A.,  
 RA Sudo H., Hosozaki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,  
 RA Niromiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,  
 RA Tanai H., Kojima M., Watanabe M., Hidaka Y., Chiba Y., Ishida S.,  
 RA Ono Y., Takiguchi S., Watanabe S., Hisada M., Hotoda T., Kusano J.,  
 RA Kaneko K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,  
 RA Togia S., Komai F., Hara F., Takeuchi K., Arita M., Imose N.,  
 RA Murashino K., Yuki H., Osima A., Sasaki N., Notsuka S.,  
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami Y., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senon A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
 RA Fujimoto Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Iitema Y., Okamoto S.,  
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senna T.,  
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togashi T., Ogawa M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Masuho Y., Yamashita R.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,

RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs";  
 RL Nat. Genet. 36:40-45 (2004).  
 DR EMBL; AK021653; BAB13865.1; -.  
 DR HSSP; Q16665; 1LQB.  
 DR GO; GO:0004871; P:signal transducer activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR001610; PAS.  
 DR Pfam; PF00987; PAS; 1.  
 DR SMART; SM00031; PAS; 2.  
 DR PROSITE; PS50112; PAS; 2.  
 DR SEQUENCE; 648 AA; 69995 MW; EBBFC74BC3F148E CRC64;  
 SQ

Query Match 35.0%; Score 577; DB 2; Length 648;  
 Best Local Similarity 61.0%; Pred. No. 5, 4e-39;  
 Matches 125; Conservative 11; Mismatches 27; Indels 42; Gaps 5;  
 QY 57 MRLTISYLHRRLCAAGGKGGRATGRLLPBPGLPVGKCO----- 109  
 Db 1 MULTISYLMRFLRGAG-----EVNQVGAGGERDADCVIKALEGP 40  
 QY 110 ---QAPGPOSVDLSSLILHNPPTGTNEF-LELIGHSIFDFIHPDCOELQALTPRN 164  
 Db 41 VMLVTAEGDMAY-----LSENVSKHGLSQLELIGHSTDFIHPDCOELQALTPQT 94  
 QY 165 LSKKCLEAPERHSLRMGSTLTSGRTNLKATWKVHLCSGHMRKPPQTSQAGP 224  
 95 LSRRKVVEAPTERCFSLRMKGSTLTSGRTNLKATWKVHLCSGHMRKPPQTSQAGP 154  
 QY 225 RSEPPLOCVLICAIPOLPFHGA 249  
 Db 155 DSEPPLOCVLICAIPOLPFHGA 249  
 SQ

SEQUENCE 199 AA; 21598 MW; 646CB066A09B47C3 CRC64;

Query Match 31.7%; Score 518.5; DB 2; Length 199;  
 Best Local Similarity 68.0%; Pred. No. 4, 5e-35;  
 Matches 104; Conservative 9; Mismatches 13; Indels 27; Gaps 2;  
 QY 120 CSSLIIRHPTPGINF-----SLELIGHSIFDFIHPDCOELQ 156  
 Db 14 CTSWLTRPSPASAPWTPRLSCASPSATCACTASARQLELIGHSTDFIHPDCOELQ 73  
 QY 157 DAUTPRMUSKLKEAPERTHSLRMGSTLTSGRTNLKATWKVHLCSGHMRKPPA 216  
 Db 74 DALTPQQTISRKVVEAPTERCFSLRMKGSTLTSGRTNLKATWKVHLCSGHMRKPPA 133  
 QY 217 QTSPPAGSRSEPPLOCVLICAIPOLPFHGA 249  
 Db 134 QTSPPAGSDSEPPLOCVLICAIPOLPFHGA 162  
 SQ

RESULT 11  
 HIF1A\_CHICK ID\_HIF1A\_CHICK STANDARD; PRT; 811 AA.  
 AC QWIB9;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DR Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha).  
 GN Name=HIF1A;  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 DE Homo sapiens (Human).  
 OC  
 NCBI\_TaxID=9031;  
 RN [1] QTSPAGSPDSEPPLOCVLICAIPOLPFHGA 162  
 RP SEQUENCE FROM N.A.  
 RC TISSUE:Heart;  
 RA Takahashi T.;  
 RT "Molecular cloning and expression of an avian cDNA for hypoxia-inducible factor-1 alpha in embryonic ventricular myocytes.";  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 CC -! FUNCTION: Functions as a master transcriptional regulator of the adaptive response to hypoxia. Binds to core DNA sequence 5'-[AGTCGCG-3', within the hypoxia response element (HRE) of target gene promoters. Activation requires recruitment of transcriptional coactivators (By similarity).  
 CC -! SUBUNIT: Efficient DNA binding requires heterodimerization of an alpha and a beta/ARNT subunit (By similarity).  
 CC -! SUBCELLULAR LOCATION: Cytoplasmic in normoxia; nuclear translocation in response to hypoxia (By similarity).  
 CC -! DOMAIN: Contains two independent C-terminal transactivation domains, NTAD and CTAD, which function synergistically. Their transcriptional activity is repressed by an intervening inhibitory domain (ID) (By similarity).  
 CC -! SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
 CC -! SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.  
 CC -! SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.  
 CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB013746; BRA34234.2; -.

DR HSSP; Q16655; IHXK.

DR InterPro; IPR001092; IHX basic.

DR InterPro; IPR001321; HypoxindIA.

DR InterPro; IPR001610; PAC.

DR InterPro; IPR000014; PAS.

DR Pfam; PF00010; IHX; 1.

DR Pfam; PF0075; PAC; 1.

DR Pfam; PF00989; PAS; 2.

DR PRINTS; PRO1080; HYPOXIAIFIA.

DR SMART; SM00353; IHX; 1.

DR SMART; SM00086; PAC; 1.

DR SMART; SM00091; PAS; 2.

DR PROSITE; PS50088; IHX; 1.

KW PROSITE; PS50112; PAS; 2.

KW Activator; DNA-binding; Nuclear protein; Repeat;

FT DNA\_BIND 17 30 Basic motif.

FT DOMAIN 31 71 Helix-loop-helix motif.

FT DOMAIN 80 157 PAS 1.

FT DOMAIN 228 298 PAS 2.

FT DOMAIN 302 345 PAC.

FT DOMAIN 401 587 ODD.

FT DOMAIN 529 573 NTAD.

FT DOMAIN 576 785 ID.

FT DOMAIN 703 706 Nuclear localization signal (Potential).

FT DOMAIN 718 721 Nuclear localization signal (Potential).

FT DOMAIN 771 811 CTAD.

FT DOMAIN 583 588 Poly-Ser.

SQ SEQUENCE 811 AA; 90542 MW; D14CD9FFC98F064CB CRC64;

Query Match 30.7%; Score 503; DB 1; Length 811;  
 Best Local Similarity 49.8%; Fred. No. 4.6e-33;  
 Matches 119; Conservative 20; Mismatches 72; Indels 28; Gaps 5;

QY 9 RNTILRKERKSDAARSRSRQQSETVYQLAHTLPFARGVSAHLKASIMRLTISYLHR 68

Db 12 RISSERKRKERSDAARCRRSKESEPYELAHQPLPHTVSAHLKASIMRLTISYLHRM 71

QY 69 LCAAGGKGKRGAT----GRUPEPGGFRGTHRRGRHGLPVFGKQQAQGPQSYDCESS 123

Db 72 LDADBELETANMEKELNCLYKALDGFVWNLSEBD-----DMIVMS 113

QY 124 LIHAPTPG-TNFSLELGHISDFTHPCOELQDALTTPRNPKKKLAPTERHPSLRM 182

Db 114 ENVKOMGLTOP--DLTGHSVDFTHPCDHELEMRLTHNGPVYKGSQNTBESFLM 171

QY 183 KSTLTSRGRTLNKATWKVHLCSCHMRAYRPPACTSPACGSPRSBPPLQCLVCEAIP 241

Db 172 KCTUTSRGRTNVIKSATWKVHLCTGHIRVDTCNQTHCG--YKKPMTCLVUCEPIP 228

RESULT 12

HIFA\_BOVIN STANDARD; PRT; 823 AA.

ID HIFA\_BOVIN STANDARD; PRT; 823 AA.

AC 0XKTA5; 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DB Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha).

GN Name=HIF1A;

OS Bos taurus (Bovine). Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; NCB\_ TaxID=9913;

[1]

DR EMBL; AB018398; BRA78675.1; -.

DR HSSP; Q16655; 1HBC.

DR InterPro; IPR001092; IHX basic.

DR InterPro; IPR001321; HypoxindIA.

DR InterPro; IPR00510; PAC.

DR InterPro; IPR000014; PAS.

DR Pfam; PF00010; IHX; 1.

DR Pfam; PF0075; PAC; 1.

DR Pfam; PF00989; PAS; 2.

DR PRINTS; PRO1080; HYPOXIAIFIA.

DR SMART; SM00353; IHX; 1.

DR SMART; SM00086; PAC; 1.

DR SMART; SM00091; PAS; 2.

RPL Biochim Biophys Acta 1445:237-243 (1999).

CC -1- FUNCTION: Functions as a master transcriptional regulator of the  
 CC adaptive response to hypoxia. Under hypoxic conditions activates  
 CC the transcription of over 40 genes, including, erythropoietin,  
 CC glucose transporters, glycolytic enzymes, vascular endothelial  
 CC growth factor, and other genes whose protein products increase  
 CC oxygen delivery or facilitate metabolic adaptation to hypoxia.  
 CC Plays an essential role in embryonic vascularization, tumor  
 CC angiogenesis and pathophysiology of ischemic disease. Binds to  
 CC core DNA sequence 5'-[AG]CTGc-3', within the hypoxia response  
 CC element (HRE) of target gene promoters. Activation requires  
 CC recruitment of transcriptional coactivators such as CREBBP and  
 CC CREBBP. Activity is enhanced by interaction with both NCOA1 or  
 CC NCOA2. Interaction with redox regulatory protein APEX seems to  
 CC activate CTAD and potentiates activation by NCOA1 and CREBBP (By  
 CC similarity).

CC -1- SUBUNIT: Bifunctional DNA binding requires heterodimerization of an  
 CC alpha and a beta/ARNT subunit. Binds to the TAZ-type 1 domains of  
 CC CREBBP and EP300. Interacts with NCOA1, NCOA2, APEX and HSP90 (By  
 CC similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic in normoxia, nuclear  
 CC translocation in response to hypoxia (By similarity).

CC -1- DOMAIN: Contains two independent C-terminal transactivation  
 CC domains, NTAD and CTAD, which function synergistically. Their  
 CC transcriptional activity is repressed by an intervening inhibitory  
 CC domain (ID) (By similarity).

CC -1- PTM: In normoxia, is hydroxylated on Pro-402 and Pro-564 in the  
 CC oxygen-dependent degradation domain (ODD) by EGLN1/PRO1 and  
 CC EGLN2/PHD2. EGLN3/PHD3 has also been shown to hydroxylate Pro-564.  
 CC The hydroxylated prolines promote interaction with VHL, initiating  
 CC rapid ubiquitination and subsequent proteasomal degradation. Under  
 CC hypoxia, proline hydroxylation is impaired and ubiquitination is  
 CC attenuated, resulting in stabilization (By similarity).

CC -1- PTM: In normoxia, is hydroxylated on Asn-800 by HIF1AN, thus  
 CC abrogating interaction with CREBBP and EP300 and preventing  
 CC transcriptional activation (By similarity).

CC -1- PTM: S-nitrosylated. All 15 free thiol groups are subjected to S-  
 CC nitrosylation in vitro, however not all thiol groups seem to be  
 CC nitrosylated in vivo (By similarity).

CC -1- PTM: Acetylation of Lys-532 by ARNT increases interaction with VHL  
 CC and stimulates subsequent proteasomal degradation (By similarity).

CC -1- PTM: Requires phosphorylation for DNA-binding (By similarity).

CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain  
 CC -1- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.

CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).

CC -----

CC DR EMBL; AB018398; BRA78675.1; -.

CC DR HSSP; Q16655; 1HBC.

CC DR InterPro; IPR001092; IHX basic.

CC DR InterPro; IPR001321; HypoxindIA.

CC DR InterPro; IPR00510; PAC.

CC DR InterPro; IPR000014; PAS.

CC DR Pfam; PF00010; IHX; 1.

CC DR Pfam; PF0075; PAC; 1.

CC DR Pfam; PF00989; PAS; 2.

CC DR PRINTS; PRO1080; HYPOXIAIFIA.

CC DR SMART; SM00353; IHX; 1.

CC DR SMART; SM00086; PAC; 1.

CC DR SMART; SM00091; PAS; 2.

DR PROSITE: PS58888; HLRH; 1.  
 DR PROSITE; PS50112; PAS; 2.  
 KW Acetylation; Activation; DNA-binding; Hydroxylation; Nuclear protein;  
 KW Phosphorylation; Repeat; S-nitrosylation; Transcription regulation.  
 FT DOMAIN 718 721 Nuclear localization signal (Potential).  
 DNA BIND 17 30 Basic motif.  
 FT DOMAIN 31 71 Helix-loop-helix motif.  
 FT DOMAIN 85 158 PAS 1.  
 FT DOMAIN 228 298 PAS 2.  
 FT DOMAIN 302 345 PAC.  
 FT DOMAIN 401 600 ODD.  
 FT DOMAIN 531 575 NTD.  
 FT DOMAIN 576 592 ID.  
 FT DOMAIN 783 823 CTD.  
 FT DOMAIN 715 718 Nuclear localization signal (Potential).  
 MOD RES 90 90 S-nitroso cysteine (Potential).  
 FT MOD-RES 139 139 S-nitroso cysteine (Potential).  
 MOD RES 173 173 S-nitroso cysteine (Potential).  
 MOD RES 194 194 S-nitroso cysteine (Potential).  
 MOD RES 210 210 S-nitroso cysteine (Potential).  
 MOD RES 219 219 S-nitroso cysteine (Potential).  
 MOD RES 224 224 S-nitroso cysteine (Potential).  
 MOD RES 255 255 S-nitroso cysteine (Potential).  
 MOD RES 334 334 S-nitroso cysteine (Potential).  
 MOD RES 337 337 S-nitroso cysteine (Potential).  
 MOD RES 359 359 Hydroxyproline (By similarity).  
 MOD RES 402 402 S-nitroso cysteine (Potential).  
 MOD RES 520 520 S-nitroso cysteine (Potential).  
 MOD RES 532 532 Hydroxyproline (By similarity).  
 MOD RES 564 564 S-nitroso cysteine (Potential).  
 MOD RES 755 755 S-nitroso cysteine (Potential).  
 MOD RES 777 777 S-nitroso cysteine (Potential).  
 MOD RES 797 797 S-nitroso cysteine (Potential).  
 MOD RES 800 800 3-hydroxyasparagine (By similarity).  
 SEQUENCE 823 AA; 92127 MW; 1267467A61B1A1 CRC64;

Query Match 30.5%; Score 499; DB 1; Length 823;  
 Best Local Similarity 46.6%; Pred. No. 1e-32; Mismatches 58; Indels 48; Gaps 6;  
 Matches 115; Conservative 26; MisMatches 58; Indels 48; Gaps 6;

QY 11 NTELKEKSRDAARSRSRQQTEVLYQLAHTLPFARGVSALDKASIMRLTISYLMRHL 70  
 14 :|:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|:  
 Db 71 AACGGRRGRATCRLIPEGPGCFRRHGTRGRGHLPGVKCQAA-----PGPQSVDLC 120  
 QY 74 DAG-----DIDIEDEKAQMNCFYLKALDGFMVLTID 105  
 Db 74 DAG-----DIDIEDEKAQMNCFYLKALDGFMVLTID 105  
 QY 121 SSSLILH----NPTPG-TNFSLFLIGHISIFDFTHPCDQOBELODALTTPRPNLSKKLEAPT 174  
 Db 106 DGDWIVYISDNVNQYMGLOF-BLGHSVPDFTHPCDNEBEMKBLTHNLNGLYRKKGQNT 163  
 QY 175 ERHFSLRMKSTLTSRGRTNLKATAWKVHLCSGHMRAYKPPQATSPGSPSREPPQCLV 234  
 Db 164 QRSFFLRLMKCTLSRGRTNLKSAWKVHLCSHMRAKPPQATSPGSPSREPPQCLV 221  
 QY 121 SSSLILH----NPTPG-TNFSLFLIGHISIFDFTHPCDQOBELODALTTPRPNLSKKLEAPT 174  
 Db 106 DGDWIVYISDNVNQYMGLOF-BLGHSVPDFTHPCDNEBEMKBLTHNLNGLYRKKGQNT 163  
 QY 235 LICEPIP 241  
 Db 222 LICEPIP 228

RESULT 13

Q61V47 PRELIMINARY; PRT; 823 AA.

ID Q61V47 05-JUL-2004 (Trembler; 27, Created)  
 DT 05-JUL-2004 (Trembler; 27, Last annotation update)  
 DT 05-JUL-2004 (Trembler; 27, Last annotation update)

DR Hypoxia inducible factor-1a.  
 GN Name=HIF-1A;  
 OS Bos mutus grunniens (Yak).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC Rattus norvegicus (Rat).  
 OS Sequence from N.A.

RC STRAIN=Wistar; TISSUE=Hepatocytes;  
 RX MEDLINE=2114367; Published=1237857; DOI=10.1042/0264-6021:3540531;  
 RA Kletzmann T.; Cornesse Y.; Brechfeld K.; Modarresi S.; Jungermann K.;  
 RT "Perivenous expression of the mRNA of the three hypoxia-inducible  
 factor a-subunits HIF-1a, HIF2a and HIF3a in rat liver.";

OC Bovinae: Bos.  
 NCBI\_TaxID=30521;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Delt K.S., Qadar Pasha M.A.;  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
 EMBL; NY62118; NAT30520\_1; --.  
 DR GO; GO:0005334; C:nucleus; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR001092; HLRH basic.  
 DR InterPro; IPR01121; HypoxindFIA.  
 DR InterPro; IPR00110; PAC.  
 DR InterPro; IPR00014; PAS.  
 Pfam; PF00010; HLRH; 1.  
 DR Pfam; PF00735; PAC; 1.  
 DR Pfam; PF00939; PAS; 2.  
 DR PRINTS; PR01080; HYPOXIAFFA.  
 DR SMART; SM00153; HLRH; 1.  
 DR SMART; SM00086; PAC; 1.  
 DR SMART; SM00931; PAS; 2.  
 DR PROSITE; PS50888; HLRH; 1.  
 DR PROSITE; PS50112; PAS; 2.  
 DR PROSITE; PS50888; HLRH; 1.  
 DR SEQUENCE 823 AA; 92128 MW; A6E388E4FEA15705 CRC64;

Query Match 30.5%; Score 499; DB 2; Length 823;  
 Best Local Similarity 46.6%; Pred. No. 1e-32; Mismatches 58; Indels 48; Gaps 6;  
 Matches 115; Conservative 26; MisMatches 58; Indels 48; Gaps 6;

QY 11 NTELKEKSRDAARSRSRQQTEVLYQLAHTLPFARGVSALDKASIMRLTISYLMRHL 70  
 14 :|:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|:  
 Db 71 AACGGRRGRATCRLIPEGPGCFRRHGTRGRGHLPGVKCQAA-----PGPQSVDLC 120  
 QY 121 SSSLILH----NPTPG-TNFSLFLIGHISIFDFTHPCDQOBELODALTTPRPNLSKKLEAPT 174  
 Db 106 DGDWIVYISDNVNQYMGLOF-BLGHSVPDFTHPCDNEBEMKBLTHNLNGLYRKKGQNT 163  
 QY 175 ERHFSLRMKSTLTSRGRTNLKATAWKVHLCSHMRAKPPQATSPGSPSREPPQCLV 234  
 Db 164 QRSFFLRLMKCTLSRGRTNLKSAWKVHLCSHMRAKPPQATSPGSPSREPPQCLV 221  
 QY 235 LICEPIP 241  
 Db 222 LICEPIP 228

RESULT 14

ID HIF1A RAT STANDARD; PRT; 825 AA.

AC 035800; Q9WU9; DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha).  
 DE Name=HIF1a;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC Rattus norvegicus (Rat).  
 OS Sequence from N.A.

RC STRAIN=Wistar; TISSUE=Hepatocytes;  
 RX MEDLINE=2114367; Published=1237857; DOI=10.1042/0264-6021:3540531;  
 RA Kletzmann T.; Cornesse Y.; Brechfeld K.; Modarresi S.; Jungermann K.;  
 RT "Perivenous expression of the mRNA of the three hypoxia-inducible  
 factor a-subunits HIF-1a, HIF2a and HIF3a in rat liver.";

RL	Biochem. J.	354:531-537 (2001).
RC	SEQUENCE FROM N.A.	[2]
RX	STRAINSprague-Dawley; TISSUE=Kidney;	DR
ZOU	MEDLINE=2147706; PubMed=11526200;	DR
RA	"oxygen-dependent expression of hypoxia-inducible factor-1alpha in renal medullary cells of rats.";	DR
RT	Physiol. Genomics 6:155-168(2001).	DR
RT	-I- FUNCTION: Functions as a master transcriptional regulator of the adaptive response to hypoxia. Under hypoxic conditions activates the transcription of over 40 genes, including, erythropoietin, glucose transporters, glycolytic enzymes, vascular endothelial growth factor, and other genes whose protein products increase oxygen delivery or facilitate metabolic adaptation to hypoxia.	DR
CC	Plays an essential role in embryonic vascularization, tumor angiogenesis and pathophysiology of ischemic disease. Binds to core DNA sequence 5'-[AG]CGTG-3', within the hypoxia response element (HRE) of target gene promoters. Activation requires recruitment of transcriptional coactivators such as CREBPP and EP300. Activity is enhanced by interaction with both, NCOAL or NCOA2. Interaction with redox regulatory protein APEX seems to activate CTAD and potentiates activation by NCOAL and CREBPP (By similarity).	DR
CC	-I- SUBUNIT: Efficient DNA binding requires heterodimerization of an alpha and a beta/ARNT subunit. Binds to the TAZ-type 1 domains of CREBPP and EP300. Interacts with NCOAL, NCOA2, APEX and Hsp90 (By similarity).	DR
CC	-I- SUBCELLULAR LOCATION: Cytoplasmic in normoxia, nuclear translocation in response to hypoxia (By similarity).	DR
CC	-I- TISSUE SPECIFICITY: Expressed in the kidney, higher expression is seen in the renal medulla than in the cortex. Expressed also in the perivenous zone of the liver.	DR
CC	-I- DOMAIN: Contains two independent C-terminal transactivation domains, NTAD and CTAD, which function synergistically. Their transcriptional activity is repressed by an intervening inhibitory domain (ID) (By similarity).	DR
CC	-I- PTM: In normoxia, is hydroxylated on Pro-402 and Pro-563 in the oxygen-dependent degradation domain (ODD) by EGLN1/PHB1 and EGLN2/PHB2. EGLN3/PHB3 has also been shown to hydroxylate Pro-563. The hydroxylated prolines promote interaction with VHL, initiating rapid ubiquitination and subsequent proteasomal degradation. Under hypoxia, proline hydroxylation is impaired and ubiquitination is attenuated, resulting in stabilization (By similarity).	DR
CC	-I- PTM: In normoxia, is hydroxylated on Asn-802 by HIF1AN, thus abrogating interaction with CREBPP and EP300 and preventing transcriptional activation (By similarity).	DR
CC	-I- PTM: S-nitrosylated. All free thiol groups are subjected to S-nitrosylation in vitro, however not all thiol groups seem to be nitrosylated in vivo (By similarity).	DR
CC	-I- PTM: Acetylation of Lys-531 by ARBD1 increases interaction with VHL and stimulates subsequent proteasomal degradation (By similarity).	DR
CC	-I- PTM: Phosphorylation is required for DNA binding (By similarity).	DR
CC	-I- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.	DR
CC	-I- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.	DR
CC	-I- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.	DR
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	DR
DR	HSSP: Q16665; 1LBC.	DR
DR	TRANSFAC: T05461; -.	DR
DR	InterPro: IPR01092; HUH basic.	DR
DR	InterPro: IPR01321; HypoxindIA.	DR
DR	InterPro: IPR01610; PAC.	DR
DR	InterPro: IPR00014; PAS.	DR
DR	EMBL: Y09507; CAA70701; -.	DR
DR	EMBL: AP057308; AAD24413; 1; -.	DR
DR	TRANSPAC: T05461; -.	DR
DR	InterPro: IPR001092; HUH basic.	DR
DR	InterPro: IPR01321; HypoxindIA.	DR
DR	InterPro: IPR01610; PAC.	DR
DR	InterPro: IPR00014; PAS.	DR
DR	PROSITE; PSS0088; HILH; 1.	DR
DR	PROSITE; PS50112; PAS; 2.	DR
KW	Acetylation; Activator; DNA-binding; Hydroxylation; Nuclear protein; Phosphorylation; Repeat; S-nitrosylation; Transcription regulation.	DR
KW	Nuclear localization signal (Potential).	DR
DOMAIN	718 721 Nuclear localization signal (Potential).	DR
DNA BIND	17 30 Basic motif.	DR
FT	MOD RES 31 71 Helix-loop-helix motif.	DR
FT	MOD RES 85 158 PAS 1.	DR
FT	MOD RES 228 298 PAS 2.	DR
FT	MOD RES 302 345 PAC.	DR
FT	MOD RES 401 602 ODD.	DR
FT	MOD RES 530 574 NTAD.	DR
FT	MOD RES 575 784 ID.	DR
FT	MOD RES 717 720 Nuclear localization signal (Potential).	DR
FT	MOD RES 785 825 CTAD.	DR
FT	MOD RES 90 90 S-nitroso cysteine (Potential).	DR
FT	MOD RES 219 219 S-nitroso cysteine (Potential).	DR
FT	MOD RES 224 224 S-nitroso cysteine (Potential).	DR
FT	MOD RES 255 334 S-nitroso cysteine (Potential).	DR
FT	MOD RES 194 194 S-nitroso cysteine (Potential).	DR
FT	MOD RES 210 210 S-nitroso cysteine (Potential).	DR
FT	MOD RES 385 385 S-nitroso cysteine (Potential).	DR
FT	MOD RES 402 402 S-nitroso cysteine (Potential).	DR
FT	MOD RES 519 519 S-nitroso cysteine (Potential).	DR
FT	MOD RES 531 531 S-nitroso cysteine (Potential).	DR
FT	MOD RES 563 563 S-nitroso cysteine (Potential).	DR
FT	MOD RES 779 779 S-nitroso cysteine (Potential).	DR
FT	MOD RES 802 802 3-hydroxyproline (By similarity).	DR
FT	MOD RES 802 802 K->NR (in Ref. 2).	DR
FT	MOD RES 12 12 D->G (in Ref. 2).	DR
FT	MOD RES 74 74 P->I (in Ref. 2).	DR
FT	MOD RES 96 96 D->N (in Ref. 2).	DR
FT	MOD RES 329 329 ATATTAT->TATA (in Ref. 2).	DR
FT	MOD RES 613 619 R->K (in Ref. 2).	DR
FT	MOD RES 708 708 R	DR
SEQUENCE	825 AA; 92319 MW; C4109AS7F38867E9 CRC64;	DR
Query	Best Local Similarity 30.4%; Score 497; DB 1; Length 825; Matches 113; Conservative 29; Mismatches 72; Indels 28; Gaps 5;	DR
QY	6 QVRVSNTLRKEKERAQDARSRSRQTEVYQOLANTLPFARGVSQNLKDASIMRLTISYL 65	DR
Db	9 EKKKNSERKKEKSDAANSRSRSKESEVPLAQMLQPLPNVNSSLDKASWMLTISYL 68	DR
QY	66 MHRLLCAG----GKRGRATGRLPEGPGFRHGTHRGRHLGLPVGKCQOAPGPOSVDLC 120	DR
Db	69 VRKULDAGDIDIEDEBMAQNCQNFVLIKAPOPGFMVLTDDG-----DMI 110	DR
QY	121 SSLSLIHNPTRG--TNSFLLEIGHSIFDFTHPCDQEQLQDALTPRENLSKKELEAPTERHFS 179	DR
Db	111 YISDNVNPKNGLTOP--ELTGHSYVFDPFDTHPCDQEMLRMLTHRGPVRKKEQENTORSFF 168	DR
QY	180 LRMSTLTSRGRTLNKAATWKVLUHSGSMRRAVEKPAQTSPPAGEPRSEPPLOCLVLECA 239	DR
Db	169 LRMKCTLTSGRTGTMKISATWKVLUHSGSMRRAVEKPAQTSPPAGEPRSEPPLOCLVLECA 226	DR
QY	240 IP 241	DR
Db	227 IP 228	DR

RESULT 15  
 Q6EMI3 PRELIMINARY; PRT; 489 AA.  
 ID Q6EMI3;  
 AC Q6EMI3;  
 DT 25-OCT-2004 (TREMBLrel. 28, Created)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)  
 DE Hypoxia-inducible factor 1 alpha (Fragment).  
 OS Xenopus laevis (African clawed frog).  
 OC Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 OC Xenopodinae; Xenopus;  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=liver;  
 RX PubMed=15162502; DOI=10.1002/dvdy.20049;  
 RA Sipe C.W., Gruber B.J., Saha M.S.;  
 RT "Short upstream region drives dynamic expression of hypoxia-inducible  
 factor 1alpha during Xenopus development.";  
 RL Dev. Dyn. 230:229-238 (2004).  
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
 DR EMBL: AV189821; AAO72733; 1; -.  
 DR GO; GO:0005634; C-nucleus; IEA.  
 DR GO; GO:0004871; Fbig1; Transcription factor activity; IEA.  
 DR GO; GO:0005355; Regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:007165; Protein transduction; IEA.  
 DR InterPro; IPR001032; HLH basic.  
 DR InterPro; IPR001331; HypoxidinIA.  
 DR InterPro; IPR001067; Nuc\_transllocat.  
 DR InterPro; IPR001610; PAC.  
 DR InterPro; IPR000014; PAS.  
 DR Pfam; PF00010; HLH; 1.  
 DR Pfam; PF000785; PAC; 1.  
 DR Pfam; PF00989; PAS; 2.  
 DR PRINTS; PR01080; HPOXKAIPIA.  
 DR SMART; SM0053; HLH; 1.  
 DR SMART; SM00086; PAC; 1.  
 DR SMART; SM00091; PAS; 2.  
 DR PROSITE; PS50888; HLH; 1.  
 DR PROSITE; PS50112; PAS; 2.  
 FT NON\_TER 1  
 SQ SEQUENCE 489 AA; 55613 MW; 3981AC3BA6AF766 CRC64;

Query Match 30.3%; Score 495; DB 2; Length 489;  
 Best Local Similarity 49.6%; Pred. No. 1.2e-32; Mismatches 71; Indels 18; Gaps 6;  
 Matches 114; Conservative 27; Mismatches 71; Indels 18; Gaps 6;

QY 13 ERKEKSRSRARSRSRQSQTVELVQLAKHTPPARGYSVAHDKASMRUTISYRMRICAA 72  
 Db 1 ERKEKSRSRARSRSRQSQTVELVQLAKHTPPARGYSVAHDKASMRUTISYRMRICAA 60

QY 73 GGKRGRATGRGLPGPGGRFRHGGRGHLGVVKCQQAPGPOSVDLGSSSLTHNPTGC- 131  
 Db 61 GDLGEGTD--LDKOLNCF---YIKALGFPLVLTREG----DMYLSEVNKCML 107

QY 132 TNFSF6LIGHISIFPFPQDQEELQDALTPRPNSKKLEAPTRHFSLRMKSTTLSRGR 191  
 Db 108 TQF--ELTGSHSVFDTHPCDHEELREMLTPRGPAKKGEQITERSFFLRMKCTVTSGR 165

QY 192 TLNUKAATWKLHSCHMAYKPPAQTSFGSPRSEPPLOCLVUCEAIP 241  
 Db 166 TNUIKSATWKLHSCHMAYKPPAQTSFGSPRSEPPLOCLVUCEAIP 213

Search completed: June 15, 2005, 16:00:09  
 Job time : 176 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

Om protein - protein search, using sw model

Run on: June 15, 2005, 16:00:19 ; Search time 157 Seconds

(Without alignments)

749.578 Million cell updates/sec

Title: US-09-896-791B-3  
Perfect score: 1636  
Sequence: 1 MAIGLQRVRSNTELRKKEKSR.....TESSLPSWVLNALNRKNCPCG 307  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters:

1710399

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cn2\_6/ptodata/1/pubpa/us08 NEW PUB pep:\*

2: /cn2\_6/ptodata/1/pubpa/pctus NEW PUB pep:\*

3: /cn2\_6/ptodata/1/pubpa/us06 PUBCOMB pep:\*

4: /cn2\_6/ptodata/1/pubpa/us07 NEW PUB pep:\*

5: /cn2\_6/ptodata/1/pubpa/pctus PUBCOMB pep:\*

6: /cn2\_6/ptodata/1/pubpa/pctus PUBCOMB pep:\*

7: /cn2\_6/ptodata/1/pubpa/us08 NEW PUB pep:\*

8: /cn2\_6/ptodata/1/pubpa/us09 NEW PUB pep:\*

9: /cn2\_6/ptodata/1/pubpa/us09C PUBCOMB pep:\*

10: /cn2\_6/ptodata/1/pubpa/us09 NEW PUB pep:\*

11: /cn2\_6/ptodata/1/pubpa/us09C PUBCOMB pep:\*

12: /cn2\_6/ptodata/1/pubpa/us10 PUBCOMB pep:\*

13: /cn2\_6/ptodata/1/pubpa/us10 PUBCOMB pep:\*

14: /cn2\_6/ptodata/1/pubpa/us10C PUBCOMB pep:\*

15: /cn2\_6/ptodata/1/pubpa/us10C PUBCOMB pep:\*

16: /cn2\_6/ptodata/1/pubpa/us10D PUBCOMB pep:\*

17: /cn2\_6/ptodata/1/pubpa/us10E PUBCOMB pep:\*

18: /cn2\_6/ptodata/1/pubpa/us10 NEW PUB pep:\*

19: /cn2\_6/ptodata/1/pubpa/us11 NEW PUB pep:\*

20: /cn2\_6/ptodata/1/pubpa/us60 NEW PUB pep:\*

21: /cn2\_6/ptodata/1/pubpa/us60 PUBCOMB pep:\*

22: /cn2\_6/ptodata/1/pubpa/us60 PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1636	100.0	307	9	US-09-896-791B-3
2	1278	78.1	324	15	US-10-307-928-20
3	795	48.6	705	14	US-10-154-386-2
4	498	30.4	827	10	US-09-910-039-149
5	498	30.4	827	14	US-10-247-671-137
6	495	30.3	623	10	US-09-967-388-2
7	495	30.3	735	16	US-10-032-561-2
8	495	30.3	826	9	US-09-922-958-4
9	495	30.3	826	9	US-09-833-79-235
10	495	30.3	826	9	US-09-736-457-330
11	495	30.3	826	9	US-09-902-941-330

### ALIGNMENTS

### RESULT 1

; Sequence 3, Application US/0996791B

; Patent No. US20020165140A1

; GENERAL INFORMATION:

; APPLICANT: Berkenstam, Anders

; APPLICANT: Bertilsson, Gran

; APPLICANT: Poellinger, Lorenz

; TITLE OF INVENTION: SCREENING METHODS

; FILE REFERENCE: 13-25-04-001

; CURRENT APPLICATION NUMBER: US/09/896,791B

; CURRENT FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: 60/217,570

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: SE 0002551-0

; PRIOR FILING DATE: 2000-07-05

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 3

; LENGTH: 307

; TYPE: PRT

; ORGANISM: Mus musculus

; US-09-896-791B-3

; Sequence 3, Appli

; Sequence 20, Appli

; Sequence 2, Appli

; Sequence 19, Appli

; Sequence 17, Appli

; Sequence 2, Appli

; Sequence 4, Appli

Query Match Similarity 100.0%; Score 1636; DB 9; Length 307;

Best Local Similarity 100.0%; Pred. No. 6; 3e-146; Mismatches 0; Indels 0; Gaps 0;

Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 330, App

Sequence 4, App

Sequence 330, App

Sequence 330, App

Sequence 4, App

Sequence 5, App

Sequence 330, App

Sequence 6, App

Sequence 7, App

Sequence 330, App

Sequence 9, App

Sequence 2, App

Sequence 330, App

Sequence 10, App

Sequence 1, App

Sequence 11, App

Sequence 12, App

Sequence 13, App

Sequence 14, App

Sequence 15, App

Sequence 16, App

Sequence 17, App

Sequence 18, App

Sequence 19, App

Sequence 20, App

Sequence 21, App

Sequence 22, App

Sequence 23, App

Sequence 24, App

Sequence 25, App

Sequence 26, App

Sequence 27, App

Sequence 28, App

Sequence 29, App

Sequence 30, App

Sequence 31, App

Sequence 32, App

Sequence 33, App

Sequence 34, App

Sequence 35, App

Sequence 36, App

Sequence 37, App

Sequence 38, App

Sequence 39, App

Sequence 40, App

Sequence 41, App

Sequence 42, App

Sequence 43, App

Sequence 44, App

Sequence 45, App

Sequence 46, App

Sequence 47, App

Sequence 48, App

Sequence 49, App

Sequence 50, App

Sequence 51, App

Sequence 52, App

Sequence 53, App

Sequence 54, App

Sequence 55, App

Sequence 56, App

Sequence 57, App

Sequence 58, App

Sequence 59, App

Sequence 60, App

Sequence 61, App

Sequence 62, App

Sequence 63, App

Sequence 64, App

Sequence 65, App

Sequence 66, App

Sequence 67, App

Sequence 68, App

Sequence 69, App

Sequence 70, App

Sequence 71, App

Sequence 72, App

Sequence 73, App

Sequence 74, App

Sequence 75, App

Sequence 76, App

Sequence 77, App

Sequence 78, App

Sequence 79, App

Sequence 80, App

Sequence 81, App

Sequence 82, App

Sequence 83, App

Sequence 84, App

Sequence 85, App

Sequence 86, App

Sequence 87, App

Sequence 88, App

Sequence 89, App

Sequence 90, App

Sequence 91, App

Sequence 92, App

Sequence 93, App

Sequence 94, App

Sequence 95, App

Sequence 96, App

Sequence 97, App

Sequence 98, App

Sequence 99, App

Sequence 100, App

Sequence 101, App

Sequence 102, App

Sequence 103, App

Sequence 104, App

Sequence 105, App

Sequence 106, App

Sequence 107, App

Sequence 108, App

Sequence 109, App

Sequence 110, App

Sequence 111, App

Sequence 112, App

Sequence 113, App

Sequence 114, App

Sequence 115, App

Sequence 116, App

Sequence 117, App

Sequence 118, App

Sequence 119, App

Sequence 120, App

Sequence 121, App

Sequence 122, App

Sequence 123, App

Sequence 124, App

Sequence 125, App

Sequence 126, App

Sequence 127, App

Sequence 128, App

Sequence 129, App

Sequence 130, App

Sequence 131, App

Sequence 132, App

Sequence 133, App

Sequence 134, App

Sequence 135, App

Sequence 136, App

Sequence 137, App

Sequence 138, App

Sequence 139, App

Sequence 140, App

Sequence 141, App

Sequence 142, App

Sequence 143, App

Sequence 144, App

Sequence 145, App

Sequence 146, App

Sequence 147, App

Sequence 148, App

Sequence 149, App

Sequence 150, App

Sequence 151, App

Sequence 152, App

Sequence 153, App

Sequence 154, App

Sequence 155, App

Sequence 156, App

Sequence 157, App

Sequence 158, App

Sequence 159, App

Sequence 160, App

Sequence 161, App

Sequence 162, App

Sequence 163, App

Sequence 164, App

Sequence 165, App

Sequence 166, App

Sequence 167, App

Sequence 168, App

Sequence 169, App

Sequence 170, App

Sequence 171, App

Sequence 172, App

Sequence 173, App

Sequence 174, App

Sequence 175, App

Sequence 176, App

Sequence 177, App

Sequence 178, App

Sequence 179, App

Sequence 180, App

Sequence 181, App

Sequence 182, App

Sequence 183, App

Sequence 184, App

Sequence 185, App

Sequence 186, App

Sequence 187, App

Sequence 188, App

Sequence 189, App

Sequence 190, App

Sequence 191, App

Sequence 192, App

Sequence 193, App

Sequence 194, App

Sequence 195, App

Sequence 196, App

Sequence 197, App

Sequence 198, App

Sequence 199, App

Sequence 200, App

Sequence 201, App

Sequence 202, App

Sequence 203, App

Sequence 204, App

Sequence 205, App

Sequence 206, App

Sequence 207, App

Sequence 208, App

Sequence 209, App

Sequence 210, App

Sequence 211, App

Sequence 212, App

Sequence 213, App

Sequence 214, App

Sequence 215, App

Sequence 216, App

Sequence 217, App

Sequence 218, App

Sequence 219, App

Sequence 220, App

Sequence 221, App

Sequence 222, App

Sequence 223, App

Sequence 224, App

Sequence 225, App

Sequence 226, App

Sequence 227, App

Sequence 228, App

Sequence 229, App

Sequence 230, App

Sequence 231, App

Sequence 232, App

Sequence 233, App

Sequence 234, App

Sequence 235, App

Sequence 236, App

Sequence 237, App

Sequence 238, App

Sequence 239, App

Sequence 240, App

Sequence 241, App

Sequence 242, App

Sequence 243, App

Sequence 244, App

Sequence 245, App

Sequence 246, App

Sequence 247, App

Sequence 248, App

Sequence 249, App

Sequence 250, App

Sequence 251, App

Sequence 252, App

Sequence 253, App

Sequence 254, App

Sequence 255, App

Sequence 256, App

Sequence 257, App

Sequence 258, App

Sequence 259, App

Sequence 260, App

Sequence 261, App

Sequence 262, App

Sequence 263, App

Sequence 264, App

Sequence 265, App

Sequence 266, App

Sequence 267, App

Sequence 268, App

Sequence 269, App

Sequence 270, App

Sequence 271, App

Sequence 272, App

Sequence 273, App

Sequence 274, App

Sequence 275, App

Sequence 276, App

Sequence 277, App

Sequence 278, App

Sequence 279, App

Sequence 280, App

Sequence 281, App

Sequence 282, App

Sequence 283, App

Sequence 284, App

Sequence 285, App

Sequence 286, App

Sequence 287, App

Sequence 288, App

Sequence 289, App

Sequence 290, App

Sequence 291, App

Sequence 292, App

Sequence 293, App

Sequence 294, App

Sequence 295, App

Sequence 296, App

Sequence 297, App

Sequence 298, App

Sequence 299, App

Sequence 300, App

Sequence 301, App

Sequence 302, App

Sequence 303, App

Sequence 304, App

Sequence 305, App

Sequence 306, App

Sequence 307, App

Sequence 308, App

Sequence 309, App

Sequence 310, App

Sequence 311, App

Sequence 312, App

Sequence 313, App

Sequence 314, App

Sequence 315, App

Sequence 316, App

Sequence 317, App

Sequence 318, App

Sequence 319, App

Sequence 320, App

Sequence 321, App

Sequence 322, App

Sequence 323, App

Sequence 324, App

Sequence 325, App

Sequence 326, App

Sequence 327, App

Sequence 328, App

Sequence 329, App

Sequence 330, App

Sequence 331, App

Sequence 332, App

Sequence 333, App

Sequence 334, App

Sequence 335, App

Sequence 336, App

Sequence 337, App

Sequence 338, App

Sequence 339, App

Sequence 340, App

Sequence 341, App

Sequence 342, App

Sequence 343, App

Sequence 344, App

Sequence 345, App

Sequence 346, App

Sequence 347, App

Sequence 348, App

Sequence 349, App

Sequence 350, App

Sequence 351, App

Sequence 352, App

Sequence 353, App

Sequence 354, App

Sequence 355, App

Sequence 356, App

Sequence 357, App

Sequence 358, App

Sequence 359, App

Sequence 360, App

Sequence 361, App

Sequence 362, App

Sequence 363, App

Sequence 364, App

Sequence 365, App

Sequence 366, App

Sequence 367, App

Sequence 368, App

Sequence 369, App

Sequence 370, App

Sequence 371, App

Sequence 372, App

Sequence 373, App

Sequence 374, App

Sequence 375, App

Sequence 376, App

Sequence 377, App

Sequence 378, App

Sequence 379, App

Sequence 380, App

Sequence 381, App

Sequence 382, App

Sequence 383, App

Sequence 384, App

Sequence 385, App

Sequence 386, App

Sequence 387, App

Sequence 388, App

Sequence 389, App

Sequence 390, App

Sequence 391, App

Sequence 392, App

Sequence 393, App

Sequence 394, App

Sequence 395, App

Sequence 396, App

Sequence 397, App

Sequence 398, App

Sequence 399, App

Sequence 400, App

Sequence 401, App

Sequence 402, App

Sequence 403, App

Sequence 404, App

Sequence 405, App

Sequence 406, App

Sequence 407, App

Sequence 408, App

Sequence 409, App

Sequence 410, App

Sequence 411, App

Sequence 412, App

Sequence 413, App

Sequence 414, App

Sequence 415, App

Sequence 416, App

Sequence 417, App

Sequence 418, App

Sequence 419, App

Sequence 420, App

Sequence 421, App

Sequence 422, App

Sequence 423, App

Sequence 424, App

Sequence 425, App

Sequence 426, App

Sequence 427, App

Sequence 428, App

Sequence 429, App

Sequence 430, App

Sequence 431, App

Sequence 432, App

Sequence 433, App

Sequence 434, App

Sequence 435, App

Sequence 436, App

Sequence 437, App

Sequence 438, App

Sequence 439, App

Sequence 440, App

Sequence 441, App

Sequence 442, App

Sequence 443, App

Sequence 444, App

Sequence 445, App

Sequence 446, App

Sequence 447, App

Sequence 448, App

Sequence 449, App

Sequence 450, App

Sequence 451, App

Sequence 452, App

Sequence 453, App

Sequence 454, App

Sequence 455, App

Sequence 456, App

Sequence 457, App

Sequence 458, App

Sequence 459, App

Sequence 460, App

Sequence 461, App

Sequence 462, App

Sequence 463, App

Sequence 464, App

Sequence 465, App

Sequence 466, App

Sequence 467, App

Sequence 468, App

Sequence 469, App

Sequence 470, App

Sequence 471, App

Sequence 472, App

Sequence 473, App

Sequence 474, App

Sequence 475, App

Sequence 476, App

Sequence 477, App

Sequence 478, App

Sequence 479, App

Sequence 480, App

Sequence 481, App

Sequence 482, App

Sequence 483, App

Sequence 484, App

Sequence 485, App

Sequence 486, App

Sequence 487, App

Sequence 488, App

Sequence 489, App

Sequence 490, App

Sequence 491, App

Sequence 492, App

Sequence 493, App

Sequence 494, App

Sequence 495, App

Sequence 496, App

Sequence 497, App

Sequence 498, App

Sequence 499, App

Sequence 500, App

Sequence 501, App

Sequence 502, App

Sequence 503, App

Sequence 504, App

Sequence 505, App

Sequence 506, App

Sequence 507, App

Sequence 508, App

Sequence 509, App

Sequence 510, App

Sequence 511, App

Sequence 512, App

Sequence 513, App

Sequence 514, App

Sequence 515, App

Sequence 516, App

Sequence 517, App

Sequence 518, App

Sequence 519, App

Sequence 520, App

Sequence 521, App

Sequence 522, App

Sequence 523, App

Sequence 524, App

Sequence 525, App

Sequence 526, App

Sequence 527, App

Sequence 528, App

Sequence 529, App

Sequence 530, App

Sequence 531, App

Sequence 532, App

Sequence 533, App

Sequence 534, App

Sequence 535, App

Sequence 536, App

Sequence 537, App

Sequence 538, App

Sequence 539, App

Sequence 540, App

Sequence 541, App

Sequence 542, App

Sequence 543, App

Sequence 544, App

Sequence 545, App

Sequence 546, App

Sequence 547, App

Sequence 548, App

Sequence 549, App

Sequence 550, App

Sequence 551, App

Sequence 552, App

Sequence 553, App

Sequence 554, App

Sequence 555, App

Sequence 556, App

Sequence 557, App

Sequence 558, App

Sequence 559, App

Sequence 560, App

Sequence 561, App

Sequence 562, App

Sequence 563, App

Sequence 564, App

Sequence 565, App

Sequence 566, App

Sequence 567, App

Sequence 568, App

Sequence 569, App

Sequence 570, App

Sequence 571, App

Sequence 572, App

Sequence 573, App

Sequence 574, App

Sequence 575, App

Sequence 5

QY 61 ISYLRMRLCAAGKGKORATGRGLPBPQPGPRHGRHGRHGLPUSKQQA PGPSVLDLC 120  
; ORGANISM: Homo sapiens  
; US-10-307-928A-20

Query Match 79.1%; Score 1278; DB 15; Length 324;  
Best Local Similarity 79.9%; Prcd. No. 4; 6e-112;  
Matches 250; Conservative 14; Mismatches 41; Indels 8; Gaps 3;

QY 181 RMKSTTSRGRTNLKAAATWKVLUHCNSCHMRAYKPPQTSPGSPRSRPLQLVLCI 240  
; 181 RMKSTTSRGRTNLKAAATWKVLUHCNSCHMRAYKPPQTSPGSPRSRPLQLVLCI 180

QY 241 POLPFHDGATGLPQECKPISTLFTPKWALKLICLVKRPVQTLQGKCTESSLPSWVIAL 300  
; 241 POLPFHDGATGLPQECKPISTLFTPKWALKLICLVKRPVQTLQGKCTESSLPSWVIAL 300

QY 301 NRKNCPG 307  
; 301 NRKNCPG 307

Db RESULT 2  
; US-10-307-928A-20  
; Sequence 20, Application US/10307928A  
; Publication No. US20030229016A1

GENERAL INFORMATION:  
; APPLICANT: Alsbrook, John P.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Boldog, Ferenc L.  
; APPLICANT: Burgess, Catherine E.  
; APPLICANT: Catterton, Elina  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Gorman, Linda  
; APPLICANT: Guo, Xiaonia (asha)  
; APPLICANT: Jil, Weizhen  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Li, Li  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Shetoy, Suresh G.  
; APPLICANT: Verner, Kimberly A.M.  
; APPLICANT: Verner, Corina A.M.  
; APPLICANT: Voss, Edward Z.

TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF USE THEREOF

FILE REFERENCE: THE SAME

CURRENT APPLICATION NUMBER: US/10/307, 928A

PRIOR FILING DATE: 2002-12-02

PRIOR APPLICATION NUMBER: 60/341,477

PRIOR FILING DATE: 2001-12-17

PRIOR APPLICATION NUMBER: 60/341,540

PRIOR FILING DATE: 2001-12-17

PRIOR APPLICATION NUMBER: 60/342,592

PRIOR FILING DATE: 2001-12-20

PRIOR APPLICATION NUMBER: 60/344,903

PRIOR FILING DATE: 2001-12-31

PRIOR APPLICATION NUMBER: 60/373, 288

PRIOR APPLICATION NUMBER: 60/380,981

PRIOR FILING DATE: 2002-05-15

PRIOR APPLICATION NUMBER: 60/381,495

PRIOR FILING DATE: 2002-05-17

PRIOR APPLICATION NUMBER: 60/383, 744

PRIOR APPLICATION NUMBER: 60/384, 024

PRIOR FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: 60/401, 788

PRIOR FILING DATE: 2002-08-07

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 53

SOFTWARE: curasedlist version 0.1

SEQ ID NO 20

LENGTH: 324

TYPE: PRT

QY 1 MAIGLQRNSNTLKERSDAARSRSRQETEYLQALTPARGVAHLKASIMRIT 60  
; 1 MAIGLQRASSTSLERKERSDAAARSRSRQETEYLQALTPARGVAHLKASIMRIT 60

QY 61 ISYLRMRLCAAGKGKORATGRGLPBPQPGPRHGRHGRHGLPUSKQQA PGPSVLDLC 120  
; 61 ISYLRMRLCAAGKGKORATGRGLPBPQPGPRHGRHGRHGLPUSKQQA PGPSVLDLC 120

QY 121 SSSLINPTPGTINFSLIGHISIFDFIHPDCOELQDALTPRPNLKCKLEAPTER 180  
; 121 SSSLINPTPGTINFSLIGHISIFDFIHPDCOELQDALTPRPNLKCKLEAPTER 180

QY 181 RMKSTTSRGRTNLKAAATWKVLUHCNSCHMRAYKPPQTSPGSPRSRPLQLVLCI 240  
; 181 RMKSTTSRGRTNLKAAATWKVLUHCNSCHMRAYKPPQTSPGSPRSRPLQLVLCI 240

QY 241 POLPFHDGATGLPQECKPISTLFTPKWALKLICLVKRPVQTLQGKCTESSLPSWVIAL 300  
; 241 POLPFHDGATGLPQECKPISTLFTPKWALKLICLVKRPVQTLQGKCTESSLPSWVIAL 300

QY 301 NRKNCPG 307  
; 301 NRKNCPG 307

Db RESULT 3  
; US-10-154-386-2  
; Sequence 2, Application US/10154386  
; Publication No. US2003026793A1

GENERAL INFORMATION:  
; APPLICANT: Angiogene Inc.  
; APPLICANT: Guy, Louis-Georges  
; TITLE OF INVENTION: HIFXIA INDUCING FACTORS AND USES THEREOF FOR INDUCING ANGIOGENESIS  
; TITLE OF INVENTION: IMPROVING MUSCULAR FUNCTIONS  
; FILE REFERENCE: 5600-81

CURRENT APPLICATION NUMBER: US/10/154,386

CURRENT FILING DATE: 2002-05-23

PRIOR APPLICATION NUMBER: US 60/292,630

PRIOR FILING DATE: 2001-05-22

PRIOR APPLICATION NUMBER: US 60/354529

PRIOR FILING DATE: 2003-02-08

NUMBER OF SEQ ID NOS: 13

SEQ ID NO 2

LENGTH: 705

TYPE: PRT

ORGANISM: Homo sapiens

US-10-154-386-2

Query Match 49.6%; Score 795; DB 14; Length 705;  
Best Local Similarity 67.6%; Prcd. No. 5; 6e-66;  
Matches 171; Conservative 11; Mismatches 29; Indels 42; Gaps 5;

QY 9 RSNTELRKSKDAARSRSRQETEVLQALHTIPFARGVSAMHDKASTMLLISYLMR 68  
; 3 RSNTELRKSKDAARSRSRQETEVLQALHTIPFARGVSAMHDKASTMLLISYLMR 62

QY 69 LCAGGKGKRATGRGLPBPQPGPRHGRHGRHGLPVGKQO-----QAPGPOSV 117  
; 69 LCAGGKGKRATGRGLPBPQPGPRHGRHGRHGLPVGKQO-----QAPGPOSV 117

Db 63 LCAGG-----EMMQVGAEGEPDACYKALEGVNMLTAEADMAY 102  
; 63 LCAGG-----EMMQVGAEGEPDACYKALEGVNMLTAEADMAY 102

QY 118 DLCSSLILHNPPTGTFNS-LEIGHISIFDFIHPDCOELQDALTPRPNLKCKLEAPTER 176  
; 103 -----LSENVSKHGLGSOLELIGHISIFDFIHPDCOELQDALTPQOTLRSRKVEAPTER 156

QY 177 HISLRMKLSTLTERGRTNLKAAATWKVLUHCNSCHMRAYKPPQTSPGSPRSRPLQLVLCI 236

RESULT 4  
US-09-919-039-149  
; Sequence 149, Application US/09919039  
; Publication No. US20030108871A1  
; GENERAL INFORMATION:  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
; FILE REFERENCE: PA-0035 US  
; CURRENT APPLICATION NUMBER: US/09/919,039  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 60/222,113  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 401  
; SOFTWARE: PERL Program  
; SEQ ID NO 149  
; LENGTH: 827  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1250434CD1  
; US-09-919-039-149  
Query Match 30.4%; Score 498; DB 10; Length 827;  
Best Local Similarity 46.2%; Pred. No. 8e-38; Mismatches 60; Indels 48; Gaps 6;  
Matches 115; Conservative 26; Mismatches 60; Indels 48; Gaps 6;  
QY 9 RSNTTELKRKRSRDAARSRSRSQETEVLYQLAHTLPARGVSAHDKASIMRLTISYLRLMR 68  
Db 13 RISSERREKRSRDAARSRSRSKESEVYELAHOLPLPHNVSSHLDKASVMRLTISYLRLRK 72  
Qy 69 LCAAGGKRGRATGRLLPGCPGGFRHGRHGRGLPVKQQA-----PGPQSVTD 118  
Db 73 LDLDG-----DLDIDDMKAQMNCYTLKALDGFWML 104  
Qy 119 LCSSLIH---NPTP-TNFSLLEIGHSIFPFDQEEQDALTPRPNLSKKEA 172  
Db 105 TDDCDMIXYSDNTYAKYMGITP-BLTGSHVDFPDRHDEKREMLTHNGLVKGKQ 162  
Qy 173 PTERHPSLRMKSTTSRGRTLNKAATWKVLHCSGHMRAYKPPAQTSAGSPRSBPPQC 232  
Db 163 NTQSFELBMKCTLSRTMTNKSATWKVLHCIGHIVDTNSNQPQCG--YKPPWTC 220  
Qy 233 LVLIICEAIP 241  
Db 221 LVLIICEPIP 229  
Db 221 LVLIICEPIP 229  
RESULT 5  
US-10-247-671-137  
; Sequence 137, Application US/10247671  
; Publication No. US20030194721A1  
; GENERAL INFORMATION:  
; APPLICANT: Mikita, Thomas  
; APPLICANT: Shiffman, Dov  
; APPLICANT: Porter, Gordon, J.  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS  
; FILE REFERENCE: PA-0050 US  
; CURRENT APPLICATION NUMBER: US/10/0247,671  
; CURRENT FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/323,784  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 186  
; SOFTWARE: PERL Program  
; SEQ ID NO 137  
; LENGTH: 827  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1250434CD1  
; US-10-247-671-137  
Query Match 30.4%; Score 498; DB 14; Length 827;  
Best Local Similarity 46.2%; Pred. No. 8e-38; Mismatches 60; Indels 48; Gaps 6;  
Matches 115; Conservative 26; Mismatches 60; Indels 48; Gaps 6;  
QY 9 RSNTTELKRKRSRDAARSRSRSQETEVLYQLAHTLPARGVSAHDKASIMRLTISYLRLMR 68  
Db 13 RISSERREKRSRDAARSRSRSKESEVYELAHOLPLPHNVSSHLDKASVMRLTISYLRLRK 72  
Qy 69 LCAAGGKRGRATGRLLPGCPGGFRHGRHGRGLPVKQQA-----PGPQSVTD 118  
Db 73 LDLDG-----DLDIDDMKAQMNCYTLKALDGFWML 104  
Qy 119 LCSSLIH---NPTP-TNFSLLEIGHSIFPFDQEEQDALTPRPNLSKKEA 172  
Db 105 TDDCDMIXYSDNTYAKYMGITP-BLTGSHVDFPDRHDEKREMLTHNGLVKGKQ 162  
Qy 173 PTERHPSLRMKSTTSRGRTLNKAATWKVLHCSGHMRAYKPPAQTSAGSPRSBPPQC 232  
Db 163 NTQSFELBMKCTLSRTMTNKSATWKVLHCIGHIVDTNSNQPQCG--YKPPWTC 220  
Qy 233 LVLIICEAIP 241  
Db 221 LVLIICEPIP 229  
Db 221 LVLIICEPIP 229  
RESULT 6  
US-09-967-388-2  
; Sequence 2, Application US/09967388  
; Publication No. US20030103956A1  
; GENERAL INFORMATION:  
; APPLICANT: JEFFERSON M. ARBEIT  
; TITLE OF INVENTION: USE OF HIF-1ALPHA VARIANTS TO ACCELERATE  
; TITLE OF INVENTION: WOUND HEALING  
; FILE REFERENCE: UC077\_001A  
; CURRENT APPLICATION NUMBER: US/09/967,388  
; CURRENT FILING DATE: 2001-09-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: HUMAN  
; US-09-967-388-2  
Query Match 30.3%; Score 495; DB 10; Length 623;  
Best Local Similarity 46.2%; Pred. No. 1.1e-37; Mismatches 59; Indels 48; Gaps 6;  
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;  
Qy 11 NTERKERSRSRDAARSRSRSQETEVLYQLAHTLPARGVSAHDKASIMRLTISYRMHUC 70  
Db 14 SSERREKRSRDAARSRSRSKESEVYELAHOLPLPHNVSSHLDKASVMRLTISYLRLK 73  
Qy 71 AAGGKRGRATGRLLPGCPGGFRHGRHGRGLPVKQQA-----PGPQSVLIC 120  
Db 74 DAG-----DLDIDDMKAQMNCYTLKALDGFWML 105  
Qy 121 SSSLIH---NPTP-TNFSLLEIGHSIFDFHPCDQOBELQDALTPRPNLSKKEAFT 174  
Db 106 DGDMIVYDYNVKMGLTQP-BLTGSHVDFPDRHDEKREMLTHNGLVKGKQNT 163  
Qy 175 ERHPSLRMKSTTSRGRTLNKAATWKVLHCSGHMRAYKPPAQTSAGSPRSBPPQCIV 234  
Db 164 QRFPLRKKCTLSRGTMNKSATWKVLHCIGHIVDTNSNQPQCG--YKPPWTCIV 221

Query Match 235 LICEAIP 241  
 Best Local Similarity 46.2%; Score 495; DB 9; Length 826;  
 US-09-922-958-4

Query Match 222 LICEPIP 228  
 Best Local Similarity 46.2%; Score 495; DB 9; Length 826;  
 US-09-922-958-4

---

RESULT 7  
 Query 11 NTBLRKEKSRDAARSRSRSQETEVLYQALHTLPFARGVSAHLDKASIMRLTISYLRMRLC 70  
 Database No. US20040214777A1  
 General Information:  
 Applicant: McGrath, Kevin  
 Title of Invention: Peptide Activators of VEGF  
 File Reference: 1443.0.020US1  
 Current Application Number: US/10/032,361  
 Current Filing Date: 2001-12-21  
 Number of SEQ ID Nos: 7  
 Software: FastSEQ for Windows Version 4.0  
 Seq ID No 2  
 Length: 735  
 Type: PRT  
 Organism: Homo sapiens  
 US-10-032-361-2

Query Match 30.3%; Score 495; DB 16; Length 735;  
 Best Local Similarity 46.2%; Pred. No. 1,3e-37; Mismatches 59; Indels 48; Gaps 6;  
 Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;  
 Sequence 2, Application US/10032361  
 Publication No. US20040214777A1

Query Match 11 NTBLRKEKSRDAARSRSRSQETEVLYQALHTLPFARGVSAHLDKASIMRLTISYLRMRLC 70  
 Database No. US20040214777A1  
 General Information:  
 Applicant: McGrath, Kevin  
 Title of Invention: Peptide Activators of VEGF  
 File Reference: 1443.0.020US1  
 Current Application Number: US/10/032,361  
 Current Filing Date: 2001-12-21  
 Number of SEQ ID Nos: 7  
 Software: FastSEQ for Windows Version 4.0  
 Seq ID No 2  
 Length: 735  
 Type: PRT  
 Organism: Homo sapiens  
 US-10-032-361-2

---

RESULT 8  
 Query Match 30.3%; Score 495; DB 16; Length 735;  
 Best Local Similarity 46.2%; Pred. No. 1,3e-37; Mismatches 59; Indels 48; Gaps 6;  
 Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;  
 Sequence 4, Application US/09922958  
 Patent No. US20020048794A1  
 General Information:  
 Applicant: Poellinger, Lorenz  
 Applicant: Pereira, Teresa  
 Applicant: Ruhn, Jorge  
 Title of Invention: MECHANISM OF CONDITIONAL REGULATION OF THE HYPOXIA-INDUCIBLE FACT  
 Title of Invention: THE VON HIPPEL-LINDAU TUMOR SUPPRESSOR PROTEIN  
 File Reference: 3743/49008  
 Current Application Number: US/09/922,958  
 Current Filing Date: 2001-08-07  
 Prior Application Number: US 60/223,480  
 Prior Filing Date: 2000-08-07  
 Number of SEQ ID Nos: 7  
 Software: PatentIn version 3.0  
 Seq ID No 4  
 Length: 826  
 Type: PRT  
 Organism: Homo sapiens  
 US-09-833-790-235

Query Match 30.3%; Score 495; DB 9; Length 826;  
 Best Local Similarity 46.2%; Pred. No. 1,5e-37; Mismatches 59; Indels 48; Gaps 6;  
 Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;  
 Sequence 4, Application US/09922958  
 Patent No. US20020048794A1  
 General Information:  
 Applicant: Poellinger, Lorenz  
 Applicant: Pereira, Teresa  
 Applicant: Ruhn, Jorge  
 Title of Invention: MECHANISM OF CONDITIONAL REGULATION OF THE HYPOXIA-INDUCIBLE FACT  
 Title of Invention: THE VON HIPPEL-LINDAU TUMOR SUPPRESSOR PROTEIN  
 File Reference: 3743/49008  
 Current Application Number: US/09/922,958  
 Current Filing Date: 2001-08-07  
 Prior Application Number: US 60/223,480  
 Prior Filing Date: 2000-08-07  
 Number of SEQ ID Nos: 7  
 Software: PatentIn version 3.0  
 Seq ID No 4  
 Length: 826  
 Type: PRT  
 Organism: Homo sapiens  
 US-09-833-790-235

---

RESULT 9  
 Query Match 30.3%; Score 495; DB 9; Length 826;  
 Best Local Similarity 46.2%; Pred. No. 1,5e-37; Mismatches 59; Indels 48; Gaps 6;  
 Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;  
 Sequence 235, Application US/09833790  
 Patent No. US20020068288A1  
 General Information:  
 Applicant: Lodes, Michael J.  
 Applicant: Wang, Tongrong  
 Applicant: Scristi, Heather  
 Applicant: Mohamath, Raodoh  
 Applicant: Fan, Liqun  
 Title of Invention: COMPOSITIONS AND METHODS FOR THE THERAPY  
 File Reference: 210121.512  
 Current Application Number: US/09/833,790  
 Current Filing Date: 2001-04-11  
 Number of SEQ ID Nos: 440  
 Software: FastSEQ for Windows Version 4.0  
 Seq ID No 235  
 Length: 826  
 Type: PRT  
 Organism: Homo sapiens  
 US-09-833-790-235

Query Match 30.3%; Score 495; DB 9; Length 826;  
 Best Local Similarity 46.2%; Pred. No. 1,5e-37; Mismatches 59; Indels 48; Gaps 6;  
 Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;  
 Sequence 235, Application US/09833790  
 Patent No. US20020068288A1  
 General Information:  
 Applicant: Lodes, Michael J.  
 Applicant: Wang, Tongrong  
 Applicant: Scristi, Heather  
 Applicant: Mohamath, Raodoh  
 Applicant: Fan, Liqun  
 Title of Invention: COMPOSITIONS AND METHODS FOR THE THERAPY  
 File Reference: 210121.512  
 Current Application Number: US/09/833,790  
 Current Filing Date: 2001-04-11  
 Number of SEQ ID Nos: 440  
 Software: FastSEQ for Windows Version 4.0  
 Seq ID No 235  
 Length: 826  
 Type: PRT  
 Organism: Homo sapiens  
 US-09-833-790-235

---

RESULT 10  
 Query Match 30.3%; Score 495; DB 9; Length 826;  
 Best Local Similarity 46.2%; Pred. No. 1,5e-37; Length 826;  
 US-09-922-958-4

US-09-736-457-330  
Sequence 330, Application US/09736457  
Patent No. US20168637A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun  
APPLICANT: Wang, Aljun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.478C15  
CURRENT APPLICATION NUMBER: US/09/736,457  
CURRENT FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 1864  
SEQ ID NO 330  
LENGTH: 826  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-736-457-330

Query Match 30.3%; Score 495; DB 9; Length 826;  
Best Local Similarity 46.2%; Pred. No. 1.5e-37;  
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

Qy 11 NTELRKEKSRDAARSRSRQETEVLYQALHTLPFARGVSAHLKDASIMRLTISYLMRHL  
Db 14 SSERKEKSRDAARSRSRKESEVFYELAHOLPLPHNVSHLKDASIMRLTISYLRVKLL 73  
Qy 71 AGKKRGRATGRLLPEPGGGFRHGHRRGRGLPVGKCOAA-----PGPOSVDLC 120  
Db 74 DAG-----DLDIEDDMKAQMCYFLKALDGFMVLTD 105  
Qy 121 SSLIH----NPPTG-TNFSLIGHISIPDFTHPCDQEELQDALTPRPNLSKCKLEAPT 174  
Db 106 DGDMDYIISDNVNKYMGITQF-BLTGHSVFDTHPCDHEEMREMLTHRNGLVKGKGEOQT 163  
Qy 175 BRHSLRMKSTLTSRGRTLNKAATWKVLUHGSGHMRAYKPPAQTSAGSPRSRPPLOCIV 234  
Db 164 QRSFFPLRMKCTLSRGRTNOKNSATWKVLUHCTGHIVHDNTSNQPCG--YKRPMPCTIV 221  
Qy 235 LICEAIP 241  
Db 222 LICEPIP 228

RESULT 12  
US-09-849-626-330  
Sequence 330, Application US/09849626  
Publication No. US202020197669A1  
GENERAL INFORMATION:  
APPLICANT: Bangur, Chaitanya  
APPLICANT: Fanger, Gary  
APPLICANT: Wang, Aljun  
APPLICANT: Wang, Tongtong  
APPLICANT: Switzer, Anne  
APPLICANT: McNeill, Patricia  
APPLICANT: Clapper, Jonathan  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.478C16  
CURRENT APPLICATION NUMBER: US/09/849,626  
NUMBER OF SEQ ID NOS: 1926  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 330  
LENGTH: 826  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-849-626-330

Query Match 30.3%; Score 495; DB 9; Length 826;  
Best Local Similarity 46.2%; Pred. No. 1.5e-37;  
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

Qy 11 NTELRKEKSRDAARSRSRQETEVLYQALHTLPFARGVSAHLKDASIMRLTISYLMRHL  
Db 14 SSERKEKSRDAARSRSRKESEVFYELAHOLPLPHNVSHLKDASIMRLTISYLRVKLL 73  
Qy 71 AGKKRGRATGRLLPEPGGGFRHGHRRGRGLPVGKCOAA-----PGPOSVDLC 120  
Db 74 DAG-----DLDIEDDMKAQMCYFLKALDGFMVLTD 105  
Qy 121 SSLIH----NPPTG-TNFSLIGHISIPDFTHPCDQEELQDALTPRPNLSKCKLEAPT 174  
Db 106 DGDMDYIISDNVNKYMGITQF-BLTGHSVFDTHPCDHEEMREMLTHRNGLVKGKGEOQT 163

QY 175 ERHFSLRMKSTLTSRGRTNLKATWKVLUHCSGHMRAYKPPAQTPAGSPRSRPPLOCIV 234  
; : ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; Db 164 QRSFLFLMKCTLSRGRTNMIKSATWKVLUHCTGHIVVDTNSNQPOCG--YKKPMPMCLV 221  
; GENERAL INFORMATION:  
; APPLICANT: JEFFERERY M. ARBEIT  
; TITLE OF INVENTION: USE OF HIF-1ALPHA VARIANTS TO ACCELERATE  
; FILE REFERENCE: UC017-001A  
; CURRENT APPLICATION NUMBER: US/09/967,388  
; CURRENT FILING DATE: 2001-09-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 826  
; TYPE: PRT  
; ORGANISM: HUMAN  
; US-09-967-388-4

RESULT 13  
; Sequence 4, Application US/09967388  
; Publication No. US2003010395A1  
; GENERAL INFORMATION:  
; APPLICANT: JEFFERERY M. ARBEIT  
; TITLE OF INVENTION: WOUND HEALING  
; FILE REFERENCE: UC017-001A  
; CURRENT APPLICATION NUMBER: US/09/967,388  
; CURRENT FILING DATE: 2001-09-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 826  
; TYPE: PRT  
; ORGANISM: HUMAN  
; US-09-967-388-4

Query Match 30.3%; Score 495; DB 10; Length 826;  
Best Local Similarity 46.2%; Pred. No. 1.5e-37; Mismatches 59; Indels 48; Gaps 6;  
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

QY 11 NTIELRKERSRDAARSRSRQSQTETVLYQALHTLPPARGVSHLDKASIMRLTISYLHRIC 70  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; Db 14 SSERRKEKSRDAARSRSRQSQTETVLYQALHTLPPARGVSHLDKASIMRLTISYLHRIC 70  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; QY 71 AAGGKGRGRATGRILLPEGGFRHGTHRRGRHGLPVGKQQA-----PGPOSVDLC 120  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; Db 74 DAG-----DLDIEDDMKAQMNCFYKLALDGFWMLTD 105  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; QY 121 SSSLIH----NPTPG-TNFSLLEIGHSIIDPFIHPCQDEEQLDALTPRPNLSKKLEAPT 174  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; Db 106 DGMIVYSDNVNKYMGTLQF--ELTGHSVFDPTHPCDHREMMLTHRLGVKGKBONT 163  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; QY 175 ERHFSLRMKSTLTSRGRTNLKATWKVLUHCSGHMRAYKPPAQTPAGSPRSRPPLOCIV 234  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; Db 164 QRSFLFLMKCTLSRGRTNMIKSATWKVLUHCTGHIVVDTNSNQPOCG--YKKPMPMCLV 221  
; CURRENT APPLICATION NUMBER: US/09/967,388  
; CURRENT FILING DATE: 2001-09-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 826  
; TYPE: PRT  
; ORGANISM: HUMAN  
; US-09-967-388-4

RESULT 15  
; Sequence 23, Application US/10028158  
; Publication No. US2002011083A1  
; GENERAL INFORMATION:  
; APPLICANT: Caniggià, Isabella  
; APPLICANT: Post, Martin  
; APPLICANT: Lye, Stephen  
; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF  
; FILE REFERENCE: 11757-38USWO  
; CURRENT APPLICATION NUMBER: US/10/028,158  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: US/09/380,662  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: PCT/CA98/00180  
; PRIOR FILING DATE: 1998-03-15  
; PRIOR APPLICATION NUMBER: US 60/039,919  
; CURRENT FILING DATE: 1997-03-07  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 23  
; LENGTH: 826  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-028-158-23

Query Match 30.3%; Score 495; DB 13; Length 826;  
Best Local Similarity 46.2%; Pred. No. 1.5e-37; Mismatches 59; Indels 48; Gaps 6;  
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

QY 11 NTIELRKERSRDAARSRSRQSQTETVLYQALHTLPPARGVSHLDKASIMRLTISYLHRIC 70  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; Db 14 SSERRKEKSRDAARSRSRQSQTETVLYQALHTLPPARGVSHLDKASIMRLTISYLHRIC 70  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; QY 71 AAGGKGRGRATGRILLPEGGFRHGTHRRGRHGLPVGKQQA-----PGPOSVDLC 120  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; Db 74 DAG-----DLDIEDDMKAQMNCFYKLALDGFWMLTD 105  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; QY 121 SSSLIH----NPTPG-TNFSLLEIGHSIIDPFIHPCQDEEQLDALTPRPNLSKKLEAPT 174  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; Db 106 DGMIVYSDNVNKYMGTLQF--ELTGHSVFDPTHPCDHREMMLTHRLGVKGKBONT 163  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; QY 175 ERHFSLRMKSTLTSRGRTNLKATWKVLUHCSGHMRAYKPPAQTPAGSPRSRPPLOCIV 234  
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
; Db 164 QRSFLFLMKCTLSRGRTNMIKSATWKVLUHCTGHIVVDTNSNQPOCG--YKKPMPMCLV 221  
; CURRENT APPLICATION NUMBER: US/09/967,388  
; CURRENT FILING DATE: 2001-09-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 826  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-028-158-23

Thu Jun 16 13:05:40 2005

usb-09-896-791b-3.rapb

Page 7

Qy 235 LICEPIP 241.  
| | |  
Db 222 LICEPIP 228

Search completed: June 15, 2005, 16:13:41  
Job time : 159 secs

This Page Blank (uspiot)

Copyright (c) 1993 - 2005 Compugen Ltd.

Om protein - protein search, using sw model

Run on: June 15, 2005, 15:50:48 ; Search time 42 Seconds

(without alignments)  
545.649 Million cell updates/sec

Title: US-09-896-791B-3

Perfect score: 1636

Sequence: 1 MALGFLQRVRSNTELRKKEKSR.....TESSLPSWVLWALNLRKNCPG 307

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/pctodata/1/iaa/5A COMB.pep: \*  
2: /cgn2\_6/pctodata/1/iaa/5B COMB.pep: \*  
3: /cgn2\_6/pctodata/1/iaa/6A COMB.pep: \*  
4: /cgn2\_6/pctodata/1/iaa/6B COMB.pep: \*  
5: /cgn2\_6/pctodata/1/iaa/PCUS COMB.pep: \*  
6: /cgn2\_6/pctodata/1/iaa/backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

\* Query

Result No. Score Match Length DB ID Description

Result No.	Score	Match	Length	DB	ID	Description
1	498	30.4	827	4	US-09-919-039-149	Sequence 149, App
2	495	30.3	245	4	US-09-438-833-3	Sequence 3, Appli
3	495	30.3	330	4	US-09-438-833-4	Sequence 4, Appli
4	495	30.3	623	4	US-09-438-838-2	Sequence 2, Appli
5	495	30.3	652	4	US-09-438-833-5	Sequence 5, Appli
6	495	30.3	813	4	US-09-438-833-12	Sequence 12, Appli
7	495	30.3	825	4	US-09-438-834-1	Sequence 6, Appli
8	495	30.3	826	2	US-09-480-2473B-2	Sequence 2, Appli
9	495	30.3	826	3	US-09-919-213-2	Sequence 2, Appli
10	495	30.3	825	3	US-09-145-547-2	Sequence 2, Appli
11	495	30.3	826	3	US-09-235-217-2	Sequence 2, Appli
12	495	30.3	826	3	US-09-380-662-23	Sequence 2, Appli
13	495	30.3	825	4	US-09-438-833-1	Sequence 1, Appli
14	495	30.3	826	4	US-09-702-705-330	Sequence 2, Appli
15	495	30.3	826	4	US-09-736-457-330	Sequence 330, App
16	495	30.3	826	4	US-09-383-581-2	Sequence 2, Appli
17	495	30.3	826	4	US-09-614-124B-330	Sequence 330, App
18	495	30.3	826	4	US-09-671-325-330	Sequence 330, App
19	495	30.3	826	4	US-09-589-184-330	Sequence 330, App
20	495	30.3	826	4	US-09-658-824-330	Sequence 330, App
21	495	30.3	826	4	US-09-959-873B-18	Sequence 18, Appli
22	495	30.3	826	4	US-09-949-016-6089	Sequence 6089, Ap
23	495	30.3	826	4	US-09-967-388-4	Sequence 4, Appli
24	495	30.3	826	5	PCT-US96-10251-2	Sequence 2, Appli
25	489	29.9	810	1	US-08-783-241-5	Sequence 7, Appli
26	477.5	29.2	875	1	US-08-785-241-5	Sequence 5, Appli
27	466	28.5	826	4	US-09-374-454-4	Sequence 4, Appli
28	455	28.5	870	1	US-08-785-241-4	Sequence 4, Appli
29	455	28.5	870	3	US-09-374-454-6	Sequence 6, Appli
30	444	27.5	805	3	US-08-910-213-4	Sequence 3, Appli
31	444	27.1	873	3	US-08-915-213-3	Sequence 3, Appli
32	444	27.1	373	3	US-09-230-217-3	Sequence 3, Appli
33	444	27.1	373	3	PCT-US96-10251-3	Sequence 3, Appli
34	444	27.1	805	2	US-08-480-473B-4	Sequence 4, Appli
35	444	27.1	805	3	US-08-910-213-4	Sequence 4, Appli
36	444	27.1	805	5	PCT-US96-10251-4	Sequence 4, Appli
37	444	27.1	805	5	PCT-US96-10251-4	Sequence 6, Appli
38	345.5	21.1	613	4	US-09-430-833-11	Sequence 11, Appli
39	345.5	19.7	756	4	US-09-919-016-7205	Sequence 7205, Ap
40	297	18.2	595	4	US-09-919-016-7205	Sequence 5, Appli
41	294	18.0	590	2	US-08-785-310B-5	Sequence 6, Appli
42	293.5	17.9	594	2	US-08-785-310B-6	Sequence 37, Appli
43	281	17.2	1507	4	US-09-914-259-37	Sequence 9161, Appli
44	263	16.1	238	4	US-09-941-016-9161	Sequence 4, Appli
45	256	15.6	570	4	US-09-923-684-4	Sequence 4, Appli

#### ALIGNMENTS

RESULT 1  
US-09-919-039-149

Sequence 149, Application US/0919039  
; Patent No. 6727066

GENERAL INFORMATION:

APPLICANT: KABER, Matthew R.  
TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES

FILE REFERENCE: PA-0035 US

CURRENT APPLICATION NUMBER: US/09/919,039

CURRENT FILING DATE: 2002-05-09

PRIOR APPLICATION NUMBER: 60/222,113

PRIOR FILING DATE: 2000-07-28

NUMBER OF SEQ ID NOS: 401

SOFTWARE: PERL Program

SEQ ID NO 149

LENGTH: 827

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6727066 1250434CD1

US-09-919-039-149

Query Match 30.4%; Score 498; DB 4; Length 827;  
Best Local Similarity 46.2%; Pred. No. 1.4e-45;  
Matches 115; Conservative 26; Mismatches 60; Indels 48; Gaps 6;

QY 9 RSNTELRLKEKSRAARSRRSQETEVLYQDQAHILPFARGVSAAHDKASTMLTISYLRRH 68

DB 13 RISSERKEKSRAKRSRKESEVFTELAHQPLPHVNSSLKDASWMLTISYLRW 72

OY 69 LCAAGGKGKRGRATGRILLPREGPGSPFRHGTRGRHGLPVCKCQDA-----PGPQSVD 118

Db 73 LLDRAG-----DLDIDDMKAQMNCFFYKLAKDGFVNL 104

QY 119 LCSSSLIH----NPTFC-TNFSLLELGHISIIPFIHDQDQEQLDALTTPRPLSKKUE 172

Db 105 TDDGDMIXTSDNTVKYMLTOF--ELTGHSVFDPTHPCDHEEMREMLTHRNGLVKKKEQ 162

QY 173 PTERRHFSIARMKSTLTSRQTLNIKAATVKJLHCSGHMRAYKUPAQTSBAGSPRSEPPQC 232

Db 163 NTQRSFELRMKCLTSRQTMKISATVKVLUCTGHTHVYDTSNQFCQG--YKKPMT 220

RESULT 2  
US-09-374-833-3

; Sequence 3, Application US/09438833  
; Patent No. 643654  
; GENERAL INFORMATION:  
; APPLICANT: Pharmacia & Upjohn  
; TITLE OF INVENTION: Protein variants  
; FILE REFERENCE: 1848  
; CURRENT APPLICATION NUMBER: US/09/438, 833  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 3  
; LENGTH: 245  
; TYPE: PRT  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain  
; OTHER INFORMATION: 1-245 of human HIF-1 alpha  
; US-09-438-833-3

Query Match 30.3%; Score 495; DB 4; Length 245;  
Best Local Similarity 46.2%; Pred. No. 4; 88-46; Mismatches 59; Indels 48; Gaps 6; Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

QY 11 NTELRKERKRSRDAARSRSRQSOETEVLYQLAHTLPPFARGVSAHLKASIMRLITISYLRMHLIC 70  
Db 14 SSERKKEKRSRDAARSRSRQSOETEVLYQLAHTLPPFARGVSAHLKASIMRLITISYLRMHLIC 70  
QY 71 AAGGKRGRATGRLLPPEGGGFRHGHRRGRHGLPVGKGCOQA-----PGPOSVDLC 120  
Db 74 DAG-----DLDIEDDMKAQNCPYLKALDGFVWLTID 105  
QY 121 SSSLIH----NPPTPG-TNFSLEIGHISIDFDIHPDCOBEQLQUALTPRNPLSKKKLEAPT 174  
Db 106 DGMIVIYISDNVNKVKGMLQTGP--BLTGHHSVDFTHPCDHREMRMLTHRNLGVKGKEONT 163  
QY 175 ERHSLRMKSTLSRGRTLNKATAWKVLUHCSGHMRAYKPPAQTSAGSPRSEPPQCLV 234  
Db 164 QRSFFLRMKCTLSRGRTLNKATAWKVLUHCSGHMRAYKPPAQTSAGSPRSEPPQCLV 221  
QY 235 LICEAIP 241  
Db 222 LICEPIP 228

RESULT 3  
US-09-438-833-4  
; Sequence 4, Application US/09438833  
; Patent No. 643654  
; GENERAL INFORMATION:  
; APPLICANT: Pharmacia & Upjohn  
; TITLE OF INVENTION: Protein variants  
; FILE REFERENCE: 1848  
; CURRENT APPLICATION NUMBER: US/09/438, 833  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 4  
; LENGTH: 330  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain  
; OTHER INFORMATION: 1-330 of human HIF-1 alpha  
; US-09-438-833-4

Query Match 30.3%; Score 495; DB 4; Length 623;  
Best Local Similarity 46.2%; Pred. No. 26-45; Mismatches 59; Indels 48; Gaps 6; Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

QY 11 NTELRKERKRSRDAARSRSRQSOETEVLYQLAHTLPPFARGVSAHLKASIMRLITISYLRMHLIC 70  
Db 14 SSERKKEKRSRDAARSRSRQSOETEVLYQLAHTLPPFARGVSAHLKASIMRLITISYLRMHLIC 70  
QY 71 AAGGKRGRATGRLLPPEGGGFRHGHRRGRHGLPVGKGCOQA-----PGPOSVDLC 120  
Db 74 DAG-----DLDIEDDMKAQNCPYLKALDGFVWLTID 105  
QY 121 SSSLIH----NPPTPG-TNFSLEIGHISIDFDIHPDCOBEQLQUALTPRNPLSKKKLEAPT 174  
Db 106 DGMIVIYISDNVNKVKGMLQTGP--BLTGHHSVDFTHPCDHREMRMLTHRNLGVKGKEONT 163  
QY 175 ERHSLRMKSTLSRGRTLNKATAWKVLUHCSGHMRAYKPPAQTSAGSPRSEPPQCLV 234  
Db 164 QRSFFLRMKCTLSRGRTLNKATAWKVLUHCSGHMRAYKPPAQTSAGSPRSEPPQCLV 221  
QY 235 LICEAIP 241  
Db 222 LICEPIP 228

RESULT 5  
US-09-438-833-5  
; Sequence 5, Application US/09438833  
; Patent No. 643654  
; GENERAL INFORMATION:  
; APPLICANT: Pharmacia & Upjohn  
; TITLE OF INVENTION: Protein variants  
; FILE REFERENCE: 1848  
; CURRENT APPLICATION NUMBER: US/09/438, 833  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1





```

; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-148-547-2

Query Match 30.3%; Score 495; DB 3; Length 826;
Best Local Similarity 46.2%; Pred. No. 3e-5;
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;
Qy 11 NTELKEKESRDAARSRSRQETEVLYQLAHTLPARGVSAHLKCASTIMLTISYLMHRLC 70
Db 14 SSERKEKSRAAARSRSRKESEVFYELAHOLPLPHNVSSHLDKASVMLTISYLRVKLL 73
Qy 71 AAGGKGKGATGRGLLPEGCGPRGTHRGRHGLPVGKQQA-----PQQSVLIC 120
Db 74 DAG-----DLDIEDDMKAQMCNCFYKLADGFVWLTID 105
Qy 121 SSSLIH----NPTPG-TNFSLEIGHSFDFIHPDCOEBLQDALTPRNLSKKLEAPT 174
Db 106 DGDMDIYISDNVNKMGLQF--ELTGHSYVDFTHPCDHEBEMRLTHRLGIVKGKGQNT 163
Qy 175 ERHFSLRMKSTLTSGRTLNKATWKVLUHCSCGMRAKYKPPAQTSPPASPRSPRBLQCLV 234
Db 164 QRSFFLRMKCTLTSRGRTNMNIKSATWKVLHCTGHIHVDTNSNQPQCG--YKKEPMTCLV 221
Qy 235 LICEAIP 241
Db 222 LICEPIP 228

RESULT 11
US-09-235-217-2
Sequence 2, Application US/09235217
Patent No. 6222018
GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HIFXOXA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/235,217
; FILING DATE: 1999-03-07
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,473
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-235-217-2

Query Match 30.3%; Score 495; DB 3; Length 826;
Best Local Similarity 46.2%; Pred. No. 3e-45;
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;
Qy 11 NTELKEKESRDAARSRSRQETEVLYQLAHTLPARGVSAHLKCASTIMLTISYLMHRLC 70
Db 14 SSERKEKSRAAARSRSRKESEVFYELAHOLPLPHNVSSHLDKASVMLTISYLRVKLL 73
Qy 71 AAGGKGKGATGRGLLPEGCGPRGTHRGRHGLPVGKQQA-----PQQSVLIC 120
Db 74 DAG-----DLDIEDDMKAQMCNCFYKLADGFVWLTID 105
Qy 121 SSSLIH----NPTPG-TNFSLEIGHSFDFIHPDCOEBLQDALTPRNLSKKLEAPT 174
Db 106 DGDMDIYISDNVNKMGLQF--ELTGHSYVDFTHPCDHEBEMRLTHRLGIVKGKGQNT 163
Qy 175 ERHFSLRMKSTLTSGRTLNKATWKVLUHCSCGMRAKYKPPAQTSPPASPRSPRBLQCLV 234
Db 164 QRSFFLRMKCTLTSRGRTNMNIKSATWKVLHCTGHIHVDTNSNQPQCG--YKKEPMTCLV 221
Qy 235 LICEAIP 241
Db 222 LICEPIP 228

RESULT 12
US-09-380-662-23
Sequence 23, Application US/09380662
Patent No. 6376199
GENERAL INFORMATION:
; APPLICANT: Caniggia, Isabella
; APPLICANT: Poet, Martin
; APPLICANT: Lye, Stephen
; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF TROPHOBLAST
; FILE REFERENCE: 11757.3B0SWO
; CURRENT APPLICATION NUMBER: US/09/380,662
; CURRENT FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: PCT/CA98/00180
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 60/039,919
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 23
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-380-662-23

Query Match 30.3%; Score 495; DB 3; Length 826;
Best Local Similarity 46.2%; Pred. No. 3e-45;
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;
Qy 11 NTELKEKESRDAARSRSRQETEVLYQLAHTLPARGVSAHLKCASTIMLTISYLMHRLC 70
Db 14 SSERKEKSRAAARSRSRKESEVFYELAHOLPLPHNVSSHLDKASVMLTISYLRVKLL 73
Qy 71 AAGGKGKGATGRGLLPEGCGPRGTHRGRHGLPVGKQQA-----PQQSVLIC 120
Db 74 DAG-----DLDIEDDMKAQMCNCFYKLADGFVWLTID 105
Qy 121 SSSLIH----NPTPG-TNFSLEIGHSFDFIHPDCOEBLQDALTPRNLSKKLEAPT 174
Db 106 DGDMDIYISDNVNKMGLQF--ELTGHSYVDFTHPCDHEBEMRLTHRLGIVKGKGQNT 163
Qy 175 ERHFSLRMKSTLTSGRTLNKATWKVLUHCSCGMRAKYKPPAQTSPPASPRSPRBLQCLV 234
Db 164 QRSFFLRMKCTLTSRGRTNMNIKSATWKVLHCTGHIHVDTNSNQPQCG--YKKEPMTCLV 221
Qy 235 LICEAIP 241
Db 222 LICEPIP 228

RESULT 13

```

US-09-438-833-1  
; Sequence 1, Application US/09438833  
; Patent No. 643654  
; GENERAL INFORMATION:  
; APPLICANT: Phamcia & Upjohn  
; TITLE OF INVENTION: Protein variant  
; FILE REFERENCE: 1848  
; CURRENT APPLICATION NUMBER: US/09/438, 833  
; CURRENT FILING DATE: 1995-11-12  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 826  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
; VOLUME: 92  
; PAGES: 5510-5514  
; DATABASE: GenBank U22431  
; DATABASE ENTRY DATE: 1995-06-28  
; US-09-438-833-1

Query Match 30.3%; Score 495; DB 4; Length 826;  
Best Local Similarity 46.2%; Pred. No. 3e-45; Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

Qy 11 NTTELKEKESRDAARSRSRQQETEVLYQALHTLPARGVS AHLKASIMRLTISYLMHRLC 70  
Db 14 SSERKEKESRDAARSRSRKESEVFYELAHQLPLPHNVSHLDKASWMLTISYLRVKKL 73

Qy 71 AAGGKGKGATGRGLPPEGGGFRGHGTHRGKHLGVPGKQQA----- pgposvnlc 120  
Db 74 DAG----- dldiedmkaoqncfyikaldgfvmltd 105

Qy 121 SSSLIH----NPTPG-TNFSLELIGHSTDFIHPQDQEBLQDALTPRNLSKKKLAAPT 174  
Db 106 DGMIVIYSDVNVKMGLQF-- ELTGHSVFDTHPCDHREMMLTHRLGTVKKGKQONT 163

Qy 175 ERHPSLRMKSTLTSRGRTNLKATWKVLUHCSCGHMRAYKPPAQTPSPASPRSEPPMQCLV 234  
Db 164 QRSFFPLRKCTLTSGRTNMNIKSATWKVLUHCIGHIHVDTNSNQPOCG-- YKPPMTCLV 221

Qy 235 LICEPIP 241  
Db 222 LICEPIP 228

RESULT 14  
; Sequence 330, Application US/09702705  
; Patent No. 6504010  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Renger, Gary  
; APPLICANT: Vedick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121-478C15  
; CURRENT APPLICATION NUMBER: US/09/736, 457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 330  
; LENGTH: 826  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-736-457-330

Query Match 30.3%; Score 495; DB 4; Length 826;  
Best Local Similarity 46.2%; Pred. No. 3e-45; Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

Qy 11 NTTELKEKESRDAARSRSRQQETEVLYQALHTLPARGVS AHLKASIMRLTISYLMHRLC 70  
Db 14 SSERKEKESRDAARSRSRKESEVFYELAHQLPLPHNVSHLDKASWMLTISYLRVKKL 73

Qy 71 AAGGKGKGATGRGLPPEGGGFRGHGTHRGKHLGVPGKQQA----- pgposvnlc 120  
Db 74 DAG----- dldiedmkaoqncfyikaldgfvmltd 105

Qy 121 SSSLIH----NPTPG-TNFSLELIGHSTDFIHPQDQEBLQDALTPRNLSKKKLAAPT 174  
Db 106 DGMIVIYSDVNVKMGLQF-- ELTGHSVFDTHPCDHREMMLTHRLGTVKKGKQONT 163

Qy 175 ERHPSLRMKSTLTSRGRTNLKATWKVLUHCSCGHMRAYKPPAQTPSPASPRSEPPMQCLV 234  
Db 164 QRSFFPLRKCTLTSGRTNMNIKSATWKVLUHCIGHIHVDTNSNQPOCG-- YKPPMTCLV 221

Thu Jun 16 13:05:39 2005

us-09-896-791b-3.rai

Page 7

Db 164 QRSFLRMKCTLSRGRTNNIKSATWKVLHCTGHIHVYDINNSNQPQCG--YKKPMTCLV 221

QY 235 LICEALP 241  
| | | |  
222 LICESPIP 228

Search completed: June 15, 2005, 16:01:42  
Job time : 43 sec<sub>B</sub>

This page blank  
ایجادی و بیکار

Copyright (c) 1993 - 2005 Compugen Ltd.

Gencore version 5.1.6

Om protein - protein search, using sw model  
Run on: June 15, 2005, 15:36:58 ; Search time 162 Seconds  
(without alignments)

US-09-896-791B-3  
Perfect score: 1636  
Sequence: MALGLQRVRSNTELRKESR.....TESSLPSWVILMALNRKNCPG 307  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues  
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseq1980s:\*
- 2: geneseq1990s:\*
- 3: geneseq2000s:\*
- 4: geneseq2001s:\*
- 5: geneseq2002s:\*
- 6: geneseq2003s:\*
- 7: geneseq2003sb:\*
- 8: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### ALIGNMENTS

RESULT 1	
AAUT75902	ID AAUT75902 standard; protein; 307 AA.
XX	XX
XX	AAUT75902;
XX	DT 08-MAY-2002 (first entry)
XX	DE Mouse inhibitory PAS domain (IPAS) protein.
XX	KW Inhibitory PAS domain protein; IPAS; cardiant; vasotropin; cardiovascular; cerebroprotective; ophthalmological; HIF-1alpha; hypoxic-induced factor 1alpha; vascular endothelial growth factor; angiogenesis; ischaemic cardiovascular lesion; stroke; diabetic microvascular disease; tumour; mouse.
XX	KW Mus musculus.
XX	OS
PN WO200202609-A1.	
XX	
PD 10-JAN-2002.	
XX	
PF 19-JUN-2001; 2001WO-SH001387.	
XX	
PR 06-JUL-2000; 2000SE-00002551.	
XX	
PA (BLOV-) BIOVITRUM AB.	
XX	
PI Berkenstam A, Bertilsson G, Poellinger L;	
XX	
DR WPI; 2002-164523/21.	
XX	
N-PSDB; ABK14502.	
XX	

US-09-896-791B-3  
Perfect score: 1636  
Sequence: MALGLQRVRSNTELRKESR.....TESSLPSWVILMALNRKNCPG 307  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseq1980s:\*
- 2: geneseq1990s:\*
- 3: geneseq2000s:\*
- 4: geneseq2001s:\*
- 5: geneseq2002s:\*
- 6: geneseq2003s:\*
- 7: geneseq2003sb:\*
- 8: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1636	100.0	307	5 AAUT5902
2	1636	100.0	307	6 ABR6227
3	1278	78.1	324	6 ABR8343
4	855.5	52.3	662	2 AAY05295
5	831	50.8	632	4 AAB93326
6	803	49.9	790	6 AAO16439
7	802	49.0	667	5 AAE4222
8	797.5	48.7	747	6 AAO16417
9	797	48.6	407	5 ABG6737
10	795	48.6	705	6 AAE4826
11	573	35.0	648	4 AAB9334
12	518.5	31.7	199	4 AAB93710
13	498	30.4	513	8 ADT16229
14	498	30.4	513	7 ADE25733
15	498	30.4	827	8 ADE6984
16	497	30.4	825	7 ADD4855
17	495	30.3	245	3 AAY94628
18	495	30.3	330	3 AAY94628
19	495	30.3	623	6 ABR0379
20	495	30.3	642	3 AAY84168
21	495	30.3	652	3 AAY94629
22	495	30.3	669	3 AAY84167
23	495	30.3	697	3 AAY84166
24	495	30.3	701	3 AAY84173
25	495	30.3	710	3 AAY84172

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseq1980s:\*
- 2: geneseq1990s:\*
- 3: geneseq2000s:\*
- 4: geneseq2001s:\*
- 5: geneseq2002s:\*
- 6: geneseq2003s:\*
- 7: geneseq2003sb:\*
- 8: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### ALIGMENTS

RESULT 1	
AAUT75902	ID AAUT75902 standard; protein; 307 AA.
XX	XX
XX	AAUT75902;
XX	DT 08-MAY-2002 (first entry)
XX	DE Mouse inhibitory PAS domain (IPAS) protein.
XX	KW Inhibitory PAS domain protein; IPAS; cardiant; vasotropin; cardiovascular; cerebroprotective; ophthalmological; HIF-1alpha; hypoxic-induced factor 1alpha; vascular endothelial growth factor; angiogenesis; ischaemic cardiovascular lesion; stroke; diabetic microvascular disease; tumour; mouse.
XX	KW Mus musculus.
XX	OS
PN WO200202609-A1.	
XX	
PD 10-JAN-2002.	
XX	
PF 19-JUN-2001; 2001WO-SH001387.	
XX	
PR 06-JUL-2000; 2000SE-00002551.	
XX	
PA (BLOV-) BIOVITRUM AB.	
XX	
PI Berkenstam A, Bertilsson G, Poellinger L;	
XX	
DR WPI; 2002-164523/21.	
XX	
N-PSDB; ABK14502.	
XX	

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseq1980s:\*
- 2: geneseq1990s:\*
- 3: geneseq2000s:\*
- 4: geneseq2001s:\*
- 5: geneseq2002s:\*
- 6: geneseq2003s:\*
- 7: geneseq2003sb:\*
- 8: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1636	100.0	307	5 AAUT5902
2	1636	100.0	307	6 ABR6227
3	1278	78.1	324	6 ABR8343
4	855.5	52.3	662	2 AAY05295
5	831	50.8	632	4 AAB93326
6	803	49.9	790	6 AAO16439
7	802	49.0	667	5 AAE4222
8	797.5	48.7	747	6 AAO16417
9	797	48.6	407	5 ABG6737
10	795	48.6	705	6 AAE4826
11	573	35.0	648	4 AAB9334
12	518.5	31.7	199	4 AAB93710
13	498	30.4	513	8 ADT16229
14	498	30.4	513	7 ADE25733
15	498	30.4	827	8 ADE6984
16	497	30.4	825	7 ADD4855
17	495	30.3	245	3 AAY94628
18	495	30.3	330	3 AAY94628
19	495	30.3	623	6 ABR0379
20	495	30.3	642	3 AAY84168
21	495	30.3	652	3 AAY94629
22	495	30.3	669	3 AAY84167
23	495	30.3	697	3 AAY84166
24	495	30.3	701	3 AAY84173
25	495	30.3	710	3 AAY84172

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseq1980s:\*
- 2: geneseq1990s:\*
- 3: geneseq2000s:\*
- 4: geneseq2001s:\*
- 5: geneseq2002s:\*
- 6: geneseq2003s:\*
- 7: geneseq2003sb:\*
- 8: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### ALIGMENTS

RESULT 1	
AAUT75902	ID AAUT75902 standard; protein; 307 AA.
XX	XX
XX	AAUT75902;
XX	DT 08-MAY-2002 (first entry)
XX	DE Mouse inhibitory PAS domain (IPAS) protein.
XX	KW Inhibitory PAS domain protein; IPAS; cardiant; vasotropin; cardiovascular; cerebroprotective; ophthalmological; HIF-1alpha; hypoxic-induced factor 1alpha; vascular endothelial growth factor; angiogenesis; ischaemic cardiovascular lesion; stroke; diabetic microvascular disease; tumour; mouse.
XX	KW Mus musculus.
XX	OS
PN WO200202609-A1.	
XX	
PD 10-JAN-2002.	
XX	
PF 19-JUN-2001; 2001WO-SH001387.	
XX	
PR 06-JUL-2000; 2000SE-00002551.	
XX	
PA (BLOV-) BIOVITRUM AB.	
XX	
PI Berkenstam A, Bertilsson G, Poellinger L;	
XX	
DR WPI; 2002-164523/21.	
XX	
N-PSDB; ABK14502.	
XX	

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseq1980s:\*
- 2: geneseq1990s:\*
- 3: geneseq2000s:\*
- 4: geneseq2001s:\*
- 5: geneseq2002s:\*
- 6: geneseq2003s:\*
- 7: geneseq2003sb:\*
- 8: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1636	100.0	307	5 AAUT5902
2	1636	100.0	307	6 ABR6227
3	1278	78.1	324	6 ABR8343
4	855.5	52.3	662	2 AAY05295
5	831	50.8	632	4 AAB93326
6	803	49.9	790	6 AAO16439
7	802	49.0	667	5 AAE4222
8	797.5	48.7	747	6 AAO16417
9	797	48.6	407	5 ABG6737
10	795	48.6	705	6 AAE4826
11	573	35.0	648	4 AAB9334
12	518.5	31.7	199	4 AAB93710
13	498	30.4	513	8 ADT16229
14	498	30.4	513	7 ADE25733
15	498	30.4	827	8 ADE6984
16	497	30.4	825	7 ADD4855
17	495	30.3	245	3 AAY94628
18	495	30.3	330	3 AAY94628
19	495	30.3	623	6 ABR0379
20	495	30.3	642	3 AAY84168
21	495	30.3	652	3 AAY94629
22	495	30.3	669	3 AAY84167
23	495	30.3	697	3 AAY84166
24	495	30.3	701	3 AAY84173
25	495	30.3	710	3 AAY84172

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseq1980s:\*
- 2: geneseq1990s:\*
- 3: geneseq2000s:\*
- 4: geneseq2001s:\*
- 5: geneseq2002s:\*
- 6: geneseq2003s:\*
- 7: geneseq2003sb:\*
- 8: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### ALIGMENTS

RESULT 1	
AAUT75902	ID AAUT75902 standard; protein; 307 AA.
XX	XX
XX	AAUT75902;
XX	DT 08-MAY-2002 (first entry)
XX	DE Mouse inhibitory PAS domain (IPAS) protein.
XX	KW Inhibitory PAS domain protein; IPAS; cardiant; vasotropin; cardiovascular; cerebroprotective; ophthalmological; HIF-1alpha; hypoxic-induced factor 1alpha; vascular endothelial growth factor; angiogenesis; ischaemic cardiovascular lesion; stroke; diabetic microvascular disease; tumour; mouse.
XX	KW Mus musculus.
XX	OS
PN WO200202609-A1.	
XX	
PD 10-JAN-2002.	
XX	
PF 19-JUN-2001; 2001WO-SH001387.	
XX	
PR 06-JUL-2000; 2000SE-00002551.	
XX	
PA (BLOV-) BIOVITRUM AB.	
XX	
PI Berkenstam A, Bertilsson G, Poellinger L;	
XX	
DR WPI; 2002-164523/21.	
XX	
N-PSDB; ABK14502.	
XX	

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Query Match 100.0%; Score 1636; DB 5; Length 307;  
 Best Local Similarity 100.0%; Pred. No. 1..9e-160; Mismatches 0; Indels 0; Gaps 0;  
 Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Sequence 307 AA;

QY 1 MALGLOVRVSNTLKRKESRDAARSRSRQSETEVLYQALANTLPFARGVSAHLKDASIMRLT 60  
 1 MALGLOVRVSNTLKRKESRDAARSRSRQSETEVLYQALANTLPFARGVSAHLKDASIMRLT 60  
 DB 61 ISYLMHRILCAAGGKRGRATGRILLPEGGFGRGTRGRHGLPVGKCOQAPGQSVLIC 120  
 61 ISYLMHRILCAAGGKRGRATGRILLPEGGFGRGTRGRHGLPVGKCOQAPGQSVLIC 120  
 QY 181 RMKSTLTSGRTNLKAATWKVLUHCSGHRAKYPAQTSPAGSPRSEBPLQCLVLCIAI 240  
 181 RMKSTLTSGRTNLKAATWKVLUHCSGHRAKYPAQTSPAGSPRSEBPLQCLVLCIAI 240  
 DB 121 SSLIHNPTPGTNSLELIGHISIDFIPCDQEEELQDALTPRPNLSKKLEAPTERHSL 180  
 121 SSLIHNPTPGTNSLELIGHISIDFIPCDQEEELQDALTPRPNLSKKLEAPTERHSL 180  
 QY 241 PQLPFHDGATGLGROEKPISTLFTPLWKALLCLVLRKWPVQVLOGKGTTESSLPSWVLWAL 300  
 241 PQLPFHDGATGLGROEKPISTLFTPLWKALLCLVLRKWPVQVLOGKGTTESSLPSWVLWAL 300  
 DB 241 PQLPFHDGATGLGROEKPISTLFTPLWKALLCLVLRKWPVQVLOGKGTTESSLPSWVLWAL 300  
 QY 301 NRKNCPG 307  
 301 NRKNCPG 307  
 DB 301 NRKNCPG 307  
 SQ Sequence 307 AA;

RESULT 2

ID ABR62227  
 ID ABR62227 standard; protein; 307 AA.

XX ABR62227;  
 XX DT 08-SEP-2003 (first entry)

DB Mouse inhibitory PAS (Per/Arnt/Sim) domain protein IPAS.

XX KW cardiant; vulnerability; cerebroprotective; antiulcer; gene therapy.

XX OS Mus sp.

XX FH Key location/Qualifiers

FT Region 14..25 /label= Basic

FT Region 26..41 /label= Helix

FT Region 42..55 /label= Loop

FT Region 56..67 /label= Helix

FT Domain 103..170 /label= PAS A

FT Domain 235..307 /label= PAS B

XX PN WO2003045440-A1.

XX PD 05-JUN-2003.

XX PR 28-NOV-2002; 2002WO-SE002198.

XX PR 28-NOV-2001; 2001WS-0333513P.

XX SQ Sequence 307 AA;

Query Match 100.0%; Score 1636; DB 6; Length 307;  
 Best Local Similarity 100.0%; Pred. No. 1..9e-160; Mismatches 0; Indels 0; Gaps 0;  
 Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Sequence 307 AA;

QY 1 MALGLOVRVSNTLKRKESRDAARSRSRQSETEVLYQALANTLPFARGVSAHLKDASIMRLT 60  
 1 MALGLOVRVSNTLKRKESRDAARSRSRQSETEVLYQALANTLPFARGVSAHLKDASIMRLT 60  
 DB 61 ISYLMHRILCAAGGKRGRATGRILLPEGGFGRGTRGRHGLPVGKCOQAPGQSVLIC 120  
 61 ISYLMHRILCAAGGKRGRATGRILLPEGGFGRGTRGRHGLPVGKCOQAPGQSVLIC 120  
 QY 121 SSLIHNPTPGTNSLELIGHISIDFIPCDQEEELQDALTPRPNLSKKLEAPTERHSL 180  
 121 SSLIHNPTPGTNSLELIGHISIDFIPCDQEEELQDALTPRPNLSKKLEAPTERHSL 180  
 QY 181 RMKSTLTSGRTNLKAATWKVLUHCSGHRAKYPAQTSPAGSPRSEBPLQCLVLCIAI 240  
 181 RMKSTLTSGRTNLKAATWKVLUHCSGHRAKYPAQTSPAGSPRSEBPLQCLVLCIAI 240  
 DB 241 PQLPFHDGATGLGROEKPISTLFTPLWKALLCLVLRKWPVQVLOGKGTTESSLPSWVLWAL 300  
 241 PQLPFHDGATGLGROEKPISTLFTPLWKALLCLVLRKWPVQVLOGKGTTESSLPSWVLWAL 300  
 QY 301 NRKNCPG 307  
 301 NRKNCPG 307  
 DB 301 NRKNCPG 307  
 SQ Sequence 307 AA;

RESULT 3

ID ABR83343  
 ID ABR83343 standard; protein; 324 AA.

XX AC ABR83343;  
 XX DT 06-OCT-2003 (first entry)

PA (ANGI-) ANGIOPHENETICS SWEDEN AB.  
 XX PI Makino Y, Cao Y, Poellinger L;  
 XX DR WPI; 2003-050158/47.  
 DR N-PSDB; ACC83867.

XX PT New pharmaceutical composition comprising IPAS, useful for increasing angiogenesis, or for treating a hypoxia-mediated condition in a cell, group of cells or organisms under hypoxic conditions, e.g. stroke or coronary heart disease.

XX Claim 4; Fig 1; 45pp; English.

The present sequence is the protein sequence of a novel murine basic helix-loop-helix-PAS (Per/Arnt/Sim) protein, termed inhibitory PAS domain protein or IPAS, that is related to hypoxia-inducible transcription factors (HIFs). IPAS has no transactivation function and demonstrates dominant negative regulation of HIF-mediated control of gene expression. Ectopic expression of IPAS in hepatoma cells selectively impaired induction of genes involved in adaptation to a hypoxic environment, and also resulted in retarded tumour growth and reduced tumor vascular density in vivo. In mice, IPAS was predominantly expressed in Purkinje cells and in corneal epithelium. Application of an IPAS antisense oligonucleotide to the mouse cornea induced angiogenesis under normoxic conditions and unmasked hypoxia-dependent induction of vascular endothelial growth factor gene expression in hypoxic cornea cells, indicating a novel role in regulation of angiogenesis and maintenance of an avascular phenotype. Claimed pharmaceutical compositions comprising IPAS antisense sequences can be used to treat increased angiogenesis, to simulate HIF-alpha function, to treat hypoxia-related conditions such as ischaemia, coronary heart disease, wound healing, stroke or diabetic ulceration, and to maintain normal cell functions under hypoxia

XX	Human NOV9a protein SEQ ID NO:20.	Db	181 RAKSTISRGRLNLKAAATWVHLGSGHRAKAYKPPAQTSPPGSPRSEPPQCLVILCAI 240
KW	anorectic; gene therapy; vaccine; cancer; neurodegenerative disorder;	Db	181 RAKSTISRGRLNLKAAATWVHLGSGHRAKAYKPPAQTSPPGSPRSEPPQCLVILCAI 240
KW	Parkinson's disease; metabolic disorder; diabetes; obesity;	Oy	241 POLPHDAGTGLPQKTPSTLFLPKL---LCLYRWPVY-LOGKGTESSLPS 294
KW	tissue typing.	Db	241 PSFLSTDGTAGTGLPQKTA-SPLSPPIYPVGRCMSLCFWRWPQIQGANGSDSLLPL 298
XX	Homo sapiens.	Db	295 WLMALNKKPG 307
XX	WO2003052061-A2.	Oy	299 EWMWALNGGNCAG 311
XX	26-JUN-2003.	Ps	03-DEC-2002; 2002WO-US038821.
XX		Ps	03-DEC-2002; 2001US-0341477P.
PR		Ps	17-DEC-2001; 2001US-0341540P.
PR		Ps	20-DEC-2001; 2001US-0342592P.
PR		Ps	31-DEC-2001; 2001US-0344903P.
PR		Ps	17-APR-2002; 2002US-0373288P.
PR		Ps	15-MAY-2002; 2002US-0380981P.
PR		Ps	17-MAY-2002; 2002US-0381495P.
DR		Ps	28-MAY-2002; 2002US-0383744P.
PR		Ps	29-MAY-2002; 2002US-0384024P.
PR		Ps	07-AUG-2002; 2002US-0401788P.
PR		Ps	26-AUG-2002; 2002US-0406353P.
PR		Ps	31-OCT-2002; 2002US-0422756P.
PR		Ps	02-DEC-2002; 2002US-040307928.
XX	(CURA-) CURAGEN CORP.	PA	
XX	Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;	PD	RESULT 4
PI	Eddinger SR, Gorman L, Guo X, Ji W, Kekuda R, Li L, Patturajan M;	ID	AY06295
PI	Rieger DK, Shenoy SG, Spytek KA, Vernet CAM, Voss EZ, Zhong M;	ID	AY06295 standard; protein; 662 AA.
XX	WPI; 2003-5333005/4.	XX	AY06295;
DR	N-PSDB; ACR06242.	AC	AY06295;
XX		XX	AY06295;
PT	New NOV9 polypeptide, useful for preparing a composition for treating or preventing e.g. cancer, neurodegenerative disorders such as diabetes or obesity, or for tissue typing.	DT	23-AUG-1999 (first entry)
PT	New NOV9 polypeptide, useful for preparing a composition for treating or preventing e.g. cancer, neurodegenerative disorders such as diabetes or obesity, or for tissue typing.	DE	Mouse transcription regulator MOP7.
PT	New NOV9 polypeptide, useful for preparing a composition for treating or preventing e.g. cancer, neurodegenerative disorders such as diabetes or obesity, or for tissue typing.	XX	MOP7; member of the PAS superfamily; bHLH-PAS; mouse; transcription regulator; hypoxia inducible factor 3 alpha.
XX		XX	transcription regulator; hypoxia inducible factor 3 alpha.
XX		OS	Mus musculus.
XX		PN	W0928464-A2.
XX		XX	W0928464-A2.
XX		PD	10-JUN-1999.
XX		XX	XX
XX		PF	AY06295;
XX		PF	27-NOV-1998; 98WO-US025314.
XX		PR	28-NOV-1997; 97US-0066863P.
XX		PA	(WISC ) WISCONSIN ALUMNI RES FOUND.
XX		XX	XX
XX		PT	Bradfield CA, Gu YZ, Hogenesch JB;
XX		XX	XX
XX		DR	WPI; 1999-371120/31.
XX		DR	N-PSDB; AAX58986.
XX		XX	XX
PT	Developmental signal transduction associated proteins.	CC	The present sequence represents mouse MOP7, a novel member of the PAS superfamily, where PAS stands for PER/ARNT/SIM domains. MOP7 cDNA (see CC AAX58986) was identified in a search of murine ESTs designed to identify CC bHLH-PAS proteins, and by RACE amplification of lung cDNA. MOP7 was CC characterised as hypoxia inducible factor 3 alpha (HIF 3 alpha). Its CC expression profile is distinct from that of HIF 1 alpha (see AAY06299), CC HIF 2 alpha (see AAY06290), MOP3 (see AAY06291). Ah receptor and Ah CC receptor nuclear translocator (ARNt), suggesting a different functional CC role. MOP7 probably regulates the same genes as HIF 1 alpha and 2 alpha, CC as evidenced by its dimerisation with the same partners (ARNt, MOP3) and CC recognition of the same core response element. MOP7 may have a functional CC role associated with response to low oxygen in the tissues in which it is CC expressed. The invention provides novel MOPs 2-9 nucleic acids (see CC AAX58981-88) and proteins (see AAY06299-97). These are useful in a CC variety of research, diagnostic and therapeutic applications. Several of CC the MOPs are alpha-class hypoxia-inducible factors. Others are involved CC in circadian signal transduction
XX	Sequence 62 AA;	SQ	
Query Match	78.1%	Score	1278; DB 6; Length 324;
Best Local Similarity	79.9%	Pred.	No. 2.5e-123; Mismatches 41; Indels 8; Gaps 3;
Matches	250;	Conservative	14; MisMatches 41; Indels 8; Gaps 3;
QY	1 MAIGLQRVRSNTELRKERSDAARSRSQETEVLYQLANTLPFARRGVWAHLDKASTMRLT 60	CC	
QY	1 MAIGLQRVRSNTELRKERSDAARSRSQETEVLYQLANTLPFARRGVWAHLDKASTMRLT 60	CC	
Db	61 ISYLMHRLLCAAGGKGKGATGRGLPPEGGGFRGTRGRHGHGIVPGKQOAPGSPVNL 120	CC	
Db	61 ISYLMHRLLCAAGGKGKGATGRGLPPEGGGFRGTRGRHGHGIVPGKQOAPGSPVNL 120	CC	
Oy	121 SSSLHNHTPGTMFSLLEGHSTFDPIHPCDQEELQDALTPRNLSKKLEAPERHSL 180	CC	
Oy	121 SSSLHNHTPGTMFSLLEGHSTFDPIHPCDQEELQDALTPRNLSKKLEAPERHSL 180	CC	
Db	121 SSSLHNHTPGTMFSLLEGHSTFDPIHPCDQEELQDALTPRNLSKKLEAPERHSL 180	CC	
Db	121 SSSLHNHTPGTMFSLLEGHSTFDPIHPCDQEELQDALTPRNLSKKLEAPERHSL 180	CC	
Oy	9 RSVNLREKRSRDRRSRSQETEVLYQLANTLPFARRGVWAHLDKASTMRLTYSRMR 68	CC	

Db 7 RSNFLRKRSKRDAAARSRSQETEVLYQLAHTLPFARGVSAHIDKASIMRLTISYLMER 66  
 CC Particularly full-length cDNAs. The primers are also useful for the  
 QY 69 LCAAGGKGRRGATGRILLPEGGPFHRGTHRGRHGLPVSKCQARGPQSVLCC----- 120  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH13642 represent human DNA sequences; AAB92446 to AAH13629 and  
 CC AAH13632 to AAH13632 represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC oligonucleotides, all of which are used in the exemplification of the  
 CC present invention.

Db 121 - -----SSSLIHNPTGTFNS-LELIGHSTIDPFIHPCDQEELQDALTPRNTLSKCL 170  
 CC

QY 95 MVLTAGEDGMAYISENVSKHIGSQLEBLIGHTSIFDFIHPCDQEELQDALTPRNTLSKCL 154  
 CC

Db 171 EAPTERHFSLKMSTLTSRGTLNITKAATWKVLUHCSGHMRAYKPAQTSPAGSPRSEPL 230  
 CC

Db 155 EAPTERHFSLKMSTLTSRGTLNITKAATWKVLUHCSGHMRAYKPAQTSPAGSPRSEPL 214  
 CC

QY 231 QCLVLICEAIP----QLPFHDGATL 251  
 CC

Db 215 QCLVLICEAIPHPASLEPPLGAFL 240  
 CC

RESULT 5

AAB93326 standard; protein; 632 AA.  
 XX

AC AAB93326;  
 XX

XX 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:12422.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

XX

EP1074617-A2.

PN

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000BP-00116126.

PR 29-JUL-1999; 999P-00248036.

AC 11-JAN-2000; 2000BP-00118775.

PR 02-MAY-2000; 2000BP-00183767.

PR 09-JUN-2000; 2000BP-00241899.

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX

PS Claim 8; SEQ ID NO 12422; 2537pp + Sequence Listing; English.

XX

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the combination of oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the combination of oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides,

CC

Db 61 ISYLMRHLCAAGKGRRGATGRILLPEGGPFHRGTHRGRHGLPVSKCQAGPQSVLCC----- 109  
 CC

Do 61 ISYLMRHLCAAG----BWNQVAGGEPDACYLKALEGVFMVL 100  
 CC

QY 110 QAGPQSVLDCSSLIHNPTGTFNS-LELIGHSTIDPFIHPCDQEELQDALTPRNTLSKCL 168  
 CC

Do 101 TAEGDMAY----LSENVIKHLGSLQELIGHTSIDFTHCQDQEELQDALTPQTLSSR 154  
 CC

QY 169 KLEAPTERHFSLKMSTLTSRGTLNITKAATWKVLUHCSGHMRAYKPAQTSPAGSPRSEPL 228  
 CC

Db 155 KLEAPTERHFSLKMSTLTSRGTLNITKAATWKVLUHCSGHMRAYKPAQTSPAGSPRSEPL 214  
 CC

QY 229 PHQCLVLICEAIPQLPFHDGA 249  
 CC

Do 215 PLOCLVLICEAIP----HPGS 231  
 CC

RESULT 6

AA016439 standard; protein; 790 AA.  
 XX

AC AA016439;  
 XX

XX 10-APR-2003 (first entry)

DE Human nucleic acid-associated protein (NAAP) - SEQ ID No 36.

XX

KW Human; nucleic acid-associated protein; NAAP; arteriosclerosis; AIDS; cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis; asthma; cancer; developmental disorder; renal tubular acidosis; anaemia; asthma; mental retardation; neurological disorder; Alzheimer's disease; epilepsy; Parkinson's disease; autoimmune disorder; inflammatory disorder; allergy; Crohn's disease; transgenic animal; animal model.

XX

OS Homo sapiens.

PN WO2003000864-A2.

XX

PD 03-JAN-2003.

XX

PF 20-JUN-2002; 2002WO-US021179.

XX

PR 22-JUN-2001; 2001US-0300518P.

PR 29-JUN-2001; 2001US-0301767P.

PR 29-JUN-2001; 2001US-0301792P.

PR 29-JUN-2001; 2001US-0301822P.

PR 29-JUN-2001; 2001US-0301833P.

PR 06-JUL-2001; 2001US-0303405P.

PR 06-JUL-2001; 2001US-0303422P.

PR 15-MAR-2002; 2002US-0364438P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

Gandhi AR, Swartakar A, Hafalia AJA, Warren BA, Emerging BM;

PI Arvizu CS, Ison CH, Honchell CD, Lee EA, Yue H, Forsythe IJ;  
 PI Ramkumar J, Griffin JA, Yang J, Sanjanawala MM, Baugh MR;  
 PT Borowsky ML, Yao MG, Walia NM, Bandhan O, Lal PG, Becha SD, Lee SY;  
 PI Richardson TW, Elliott VS, Luo W, Tang YR, Zebanjian Y, Lu Y;  
 XX DR WPI; 2003-201420/19.  
 DR N-PSDB; AAL51589.

XX PT New nucleic acid-associated proteins and polynucleotides, useful for  
 PT diagnosing, treating or preventing cell proliferative (e.g. cancer),  
 PT neurological (e.g. epilepsy or Parkinson's disease), or autoimmune  
 PT disorders (e.g. AIDS).

XX PS Claim 1; Page 274-276; 312pp; English.

CC The invention comprises the amino acid and coding sequences of human  
 CC nucleic acid-associated proteins (NAAP). The DNA and protein sequences of  
 CC the invention are useful for diagnosing, treating or preventing disorders  
 CC associated with aberrant expression of NAAP, such as: cell proliferative  
 CC disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis,  
 CC or mental retardation); neurological disorders (e.g. Alzheimer's disease,  
 CC Parkinson's disease or epilepsy); and autoimmune/inflammatory disorders  
 CC (e.g. AIDS, allergies, asthma) or Crohn's disease. The DNA sequences of  
 CC the invention are useful for creating transgenic animals to model human  
 CC disease. The present amino acid sequence represents a human nucleic acid-  
 XX SQ Sequence 790 AA;

Query Match 49.1%; Score 803; DB 6; Length 790;  
 Best Local Similarity 58.7%; Prod. No. 1.5e-73; Mismatches 179; Indels 11; Gaps 6;  
 Matches 179; Conservative 11; Mismatches 29; Indels 86; Gaps 6;

QY 1 MALGIQRVR-----SNTELRK 16  
 1 MALGIQRARALSCCVISPPCAPTRNSHRCGGCTASPPAPPGWPSQRPGRMSTELRK 60  
 17 EKSDAARSRSRSQETEVLYQLAHTLPFARGVSVAHLDKASIMRLTISYLRLCAAGGER 76  
 61 EKSDAARSRSRSQETEVLYQLAHTLPFARGVSVAHLDKASIMRLTISYLRLCAAG--- 117  
 77 GRATGRLLPBGPGGTRRHGRHGLPYGKQ-----QAPGQSVQDCLSSL 125  
 118 -----EMNQVAGGEPLDACYLKALEGFVFMVLTAEQDMAY-----LS 154

Db QY 126 HNPPTGTTNFS-LELGHSTPDFIHCQDQEQLDALTPTRNLSKKLEAPERHRFLRMS 184  
 155 ENVSKHLGSQLELIGHSTPDFIHCQDQEQLDALTPTQQTLSRKVEATERCESRMS 214

Db QY 185 TLTSRGRTNLKAATWKVLCSHGRAYKPAQTPSPAGSPRSPERPLQCLVLTCAIPOLP 244  
 215 TLTSRGRTNLKAATWKVLCSHGRAYKPAQTPSPAGSPDSEPPLOCLVLTCAIP--- 271

QY 245 FHDGA 249  
 Db 272 -HPGS 275

, RESULT 7  
 AAB24222  
 ID AAB24222 standard; protein; 667 AA.  
 XX AC AAB24222;  
 XX DT 23-SEP-2002 (first entry)  
 XX DB Human HIF-3 protein.

KW angiogenic myeloid metaplasia; myeloid leukaemia; gene therapy;  
 KW polycythaemia vera; hypoxia responsive element; HRE.  
 XX Homo sapiens.  
 OS WO200231291-12.  
 XX PN 02-MAY-2002.  
 XX PR 25-OCT-2001; 2001WO-US049856.  
 XX PR 26-OCT-2000; 2000US-0243542P.  
 XX PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
 XX PI Colgan SP;  
 DR N-PSDB; AAD39042.  
 XX WPI; 2002-471427/50.

CC Treating a subject (at risk of) having a hematologic malignancy or  
 CC multidrug resistance, e.g. lymphoma or myeloma, by administering hypoxia  
 CC inducible factor-1 binding molecules or small ubiquitin-like-modifier-1  
 PT binding molecules.  
 XX Disclosure; Page 88-91; 92pp; English.

CC The invention relates to a method of treating a subject having or at risk  
 CC of developing a hematologic malignancy or multidrug resistance (MDR).  
 The method involves administering hypoxia inducible factor-1 (HIF-1)  
 CC binding molecules or small ubiquitin-like-modifier (SMO)-1 binding  
 molecules or HIF-1-SMO-1 complex modulators. mdr-hypoxia responsive  
 CC element (HRE) binding molecules or antisense nucleic acid molecules and  
 CC SMO-1 binding molecules or antisense molecules are useful for treating a  
 CC subject having or at risk of developing hematologic malignancy or MDR  
 CC (e.g. a lymphoid disorder or a myeloid disorder). The lymphoid disorders  
 CC include lymphocytic leukaemia or chronic lymphoproliferative disorders  
 e.g. Lymphoma, myeloma or chronic lymphoid Leukaemia. The myeloid  
 CC disorders include chronic or acute myeloid leukaemia, e.g. angiogenic  
 CC myeloid metaplasia, essential thrombocythaemia or polycythaemia vera. The  
 CC invention is used in gene therapy. The present sequence is human HIF-3  
 XX protein  
 SQ Sequence 667 AA;

Query Match 49.0%; Score 802; DB 5; Length 667;  
 Best Local Similarity 68.0%; Prod. No. 1.5e-72; Mismatches 172; Indels 42; Gaps 5;  
 Matches 172; Conservative 11; Mismatches 28; Indels 42; Gaps 5;

QY 9 RSNTELRKESRDAARSRSQETEVLYQLAHTLPFARGVSVAHLDKASIMRLTISYLRL 68  
 7 RSTTELRKESRDAARSRSQETEVLYQLAHTLPFARGVSVAHLDKASIMRLTISYLRL 66  
 QY 69 LCAAGKGKRGRATGRLLPHPGPGGTRRHGRHGLPYGKQ-----QAPGQSV 117  
 67 LCAGG-----EMNQVAGGEPLDACYLKALEGFVFMVLTAEQDMAY 106

Db QY 118 DICSSSLIHNPPTGTTNFS-LELGHSTPDFIHCQDQEQLDALTPTRNLSKKLEAPER 176  
 107 -----EMNQVAGGEPLDACYLKALEGFVFMVLTAEQDMAY 160

Db QY 177 HPSLRMKSTLTSRGRTNLKAATWKVLCSHGRAYKPAQTPSPAGSPRSPERPLQCLVLI 236  
 161 CFSLRMKSTLTSRGRTNLKAATWKVLCSHGRAYKPAQTPSPAGSPDSEPPLOCLVLI 220

QY 237 CEALPQLDFHDGA 249  
 Db 221 CEAIP---HPGS 229

XX  
AC AAO16417;  
XX  
DT 10-APR-2003 (first entry)  
XX  
DE Human nucleic acid-associated protein (NAAP) - SEQ ID No 14.  
XX  
KW Human; nucleic acid-associated protein; NAAP; arteriosclerosis; cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis; AIDS; cancer; developmental disorder; renal tubular acidosis; anaemia; asthma; mental retardation; neurological disorder; Alzheimer's disease; epilepsy; Parkinson's disease; autoimmune disorder; inflammatory disorder; allergy; Crohn's disease; transgenic animal; animal model.  
XX  
OS Homo sapiens.  
XX  
PN WO2003000864-A2.  
XX  
PD 03-JAN-2003.  
XX  
PP 20-JUN-2002; 2002WO-US021179.  
XX  
PR 22-JUN-2001; 2001US-0300518P.  
PR 29-JUN-2001; 2001US-0301787P.  
PR 29-JUN-2001; 2001US-0301792P.  
PR 29-JUN-2001; 2001US-0301892P.  
PR 29-JUN-2001; 2001US-0301893P.  
PR 06-JUL-2001; 2001US-0303405P.  
PR 15-MAR-2002; 2002US-0364438P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PT Gandhi AR, Swarnkar A, Hafalia AJA, Warren BA, Emerling BM; PI Arvizu CS, Ison CH, Honchell CD, Lee EA, Yue H, Foreythe IJ; PI Ramkumar J, Griffin JA, Yang J, Sanjanwala MM, Baughn MR; PI Borowsky MI, Yao MG, Walla NK, Bandman O, Iai PG, Becha SD, Lee SY; PT Richardson TW, Elliott VS, Luo W, Tang YT, Zabarjadian Y, Lu Y; DR WPI; 2003-201420/19.  
XX  
DR N-PSDB; ABK1567.  
XX  
PT New nucleic acid-associated proteins and polymucleotides, useful for diagnosing, treating or preventing cell proliferative (e.g. cancer), neurological (e.g. epilepsy or Parkinson's disease), or autoimmune disorders (e.g. AIDS).  
XX  
PS Claim 1; Page 232-234; 312pp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of human nucleic acid-associated proteins (NAAP). The DNA and protein sequences of the invention are useful for diagnosing, treating or preventing disorders associated with aberrant expression of NAAP, such as: cell proliferative disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis or cancer); developmental disorders (e.g. renal tubular acidosis, anaemia or mental retardation); neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or epilepsy); and autoimmune/inflammatory disorders (e.g. AIDS, allergies, asthma or Crohn's disease). The DNA sequences of the invention are useful for creating transgenic animals to model human disease. The present amino acid sequence represents a human nucleic acid-associated protein of the invention  
XX  
SQ Sequence 747 AA;  
Query Match 48.7%; Score 797.5; DB 6; Length 747;  
Best Local Similarity 60.4%; Pred. No. 5.2e-73;  
Matches 180; Conservative 10; Mismatches 35; Indels 73; Gaps 6;  
QY 1 MALGQVRV-----SNTELRK 16  
Db 1 MALGQARPALSCCVISPPCAPTRNSHPGCTASPPAPPWPFSSQRGPGRNSTELRK 60  
QY 17 EKSRDAAKRSRSQEETIVLQVLAHTTPAKGSVSAHKASIMRITISVLMRCAAGKK 76

DB 61 EKSRDAAKRSRSQEETIVLQVLAHTTPAKGSVSAHKASIMRITISVLMRCAAGKK 120  
QY 77 GRATGR---LIPPEGPGCFRHRGRHGLPVGKCOQPGPOSVDUCCSSLHNPTPGT 132  
DB 121 QVGAGENWWMLIKALESGFVWLTAGG-----DM--AVISENTSKHL 160  
QY 133 NFS-LELIGHSIFDPTRPCDQERLQALTTPRNLSKCKLEAPTERHFSLRMKSTLSRGR 191  
Db 161 GLSQLELIGHSFDFIFHCDCDQBLLQALTTPQQLTSRKVEAPTERCFSLRMKSTLSRGR 220  
QY 192 TNLKAATWKVLUHCSGMRAYKPPQTSPAGSPRSRSPPLQCLVLCIAIPOLPHOGA 249  
DB 221 TNLKAATWKVLUHCSGMRAYKPPQTSPAGSPRSRSPPLQCLVLCIAIP---HPGS 274

RESULT 9  
ABG66737  
ID ABG66737 standard, protein: 407 AA.  
XX  
AC ABG66737;  
XX  
DT 30-AUG-2002 (first entry)  
XX  
DE Human novel polypeptide #72.  
XX  
KW Human; inflammatory condition; shock; sepsis; immune response; cancer; wound healing; central nervous system disease; haematoopoiesis; peripheral nervous system disease; amyotrophic lateral sclerosis; tendon; myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone; cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis; bone degenerative disorder; periodontal disease; reperfusion injury; lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection; allergic condition; thrombolytic; thrombosis; coagulation disorder; fungal infection.  
XX  
OS Homo sapiens.  
XX  
PN WO2002A4340-A2.  
XX  
PD 06-JUN-2002.  
XX  
PP 30-NOV-2001; 2001WO-US047004.  
XX  
PR 30-NOV-2000; 2000US-00728952.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PT Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D; PI Yamazaki V, Ujjwal ML, Drmanac RT; DR WPI; 2002-508509/54.  
N-PSDB; ABK94961.  
XX  
PT Novel nucleic acids and polypeptides for diagnosis, treatment of inflammatory, autoimmune, nervous system, myeloid or lymphoid cell disorders, cancer and promoting wound healing.  
XX  
PS Claim 10; Page 649; 672pp; English.  
XX  
CC The invention relates to human novel polymucleotides and associated polypeptides. The polymucleotides and polypeptides are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses and cancer and for promoting wound healing. The sequences are used to induce the proliferation of neural cells and regeneration of nerve and brain tissue, and are useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's disease, Parkinson's disease, Huntington's disease and amyotrophic lateral sclerosis. The sequences are involved in chemotactic or chemokinetic activity, regulation of haemopoiesis, treatment of myeloid or lymphoid cell disorders and platelet disorders, such as thrombocytopenia, regeneration of bone, cartilage, tendon, ligament and/or nerve tissue.

CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of

CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal

CC disease. The sequences of the invention are also useful for gut

CC protection or regeneration and treatment of lung or liver fibrosis,

CC reperfusion injury in various tissues, immune deficiencies and disorders

CC including severe combined immunodeficiency (SCID), bacterial or fungal

CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia

CC gravis, allergic conditions such as asthma, thrombolytic or thrombosis

CC and coagulation disorders. Sequences ABG66666-ABC66758 represent human

CC novel polypeptides of the invention

XX SQ Sequence 407 AA;

Query Match 48.7%; Score 797; DB 5; Length 407;

Best Local Similarity 67.9%; Pred. No. 2.4e-73; Mismatches 171; Conservatve 11; Indels 42; Gaps 5;

Db 10 SNTELRKERKSRAARSRSRQQETEVLYQLAHTLPFARGVS AHLDKASIMRLTISYLRMHL 69

Qy 54 SNTTELRKERKSRAARSRSRQQETEVLYQLAHTLPFARGVS AHLDKASIMRLTISYLRMHL 113

Db 70 CAAGSKKRGRATGRILLPGEGPGFRHGRHRGLGVKGCO-----QAGPQSTD 118

Db 114 CAAG-----EWNQVAGGEEPLACYLKALEGFVWMLTAEGDMAY- 152

Qy 119 LCSSSLIHNPTPGTNP-LELIGHISIPDFIHCDOELQDALTRPNLUSKKLAEAPERH 177

Db 153 ---LSENVSKHIGLSQELIGHISIFDFTHCDQEELQDALTPQQTLSRRKVAPTERC 207

Qy 178 FSLRMKSTLTSRGRTLNKAATWKVLUHCSCMHMRAYKPAOTS PAGSPRSEPPQCLVLIC 237

Db 208 FSLRMKSTLTSRGRTLNKAATWKVLUHCSCMHMRAYKPAOTS PAGSPDSEPLQCLVLIC 267

Qy 238 EAIPOLPFHDGA 249

Db 268 EAIP---HPGS 275

RESULT 10

AAB34826 ID AAB34826 standard; protein; 705 AA.

AC AAB34826; XX DT 28-MAY-2003 (first entry)

DB Human hypoxia inducible factor (HIF)- $\alpha$  protein.

XX Human; hypoxia inducible factor; HIF- $\alpha$ ; tumour; muscular function; vascular endothelial growth factor; VEGF; angiogenesis; cardiovascular; coronary disease; cardiac disease; cytostatic; Cardiant; cell survival; OS Homo sapiens.

XX WO200294862-A2. PN

XX PD 28-NOV-2002.

XX PR 23-MAY-2002; 2002WO-CA000752.

XX PR 23-MAY-2001; 2001NS-0292630P.

XX PR 08-FEB-2002; 2002US-0354529P.

PA (ANGI-) ANGIOGENE INC.

XX PT Guy L; PI

DR WPI: 2003-129410/12.

DR N-FSDB; AD53131.

XX PT New Hypoxia Inducible Factor-3 alpha polynucleotide for inducing vascular endothelial growth factor (VEGF) expressions in a tissue or cell, particularly for eliminating tumors, or treating coronary or cardiac

PT diseases in mammals.

XX PS Claim 13; Col 69-72; 38pp; English.

The invention relates to hypoxia inducible factor (HIF)- $\alpha$  DNA and protein sequences. The HIF- $\alpha$  DNA and protein sequences are useful for inducing vascular endothelial growth factor (VEGF) expression or angiogenesis in a mammalian tissue or cell, for modulating tumoural cell survival, eliminating a tumoural cell, or evaluating the malignancy of a tumour in a subject. They are also useful for improving muscular function, particularly for treating coronary and cardiac diseases in

CC mammals. The present sequence is human HIF- $\alpha$  protein.

XX SQ Sequence 705 AA;

Query Match 48.6%; Score 795; DB 6; Length 705;

Best Local Similarity 67.6%; Pred. No. 8.6e-73; Mismatches 171; Conservatve 11; Indels 42; Gaps 5;

Db 9 RSNTELRKERKSRAARSRSRQQETEVLYQLAHTLPFARGVS AHLDKASIMRLTISYLRMHL 68

Qy 3 RSTTELRKERKSRAARSRSRQQETEVLYQLAHTLPFARGVS AHLDKASIMRLTISYLRMHL 62

Db 69 LCAGGKGRERATGRILLPGEGPGFRHGRHRGLPVSKCQ-----QAGPQSV 117

Qy 63 LCAG-----EWNQVAGGEEPLACYLKALEGFVWMLTAEGDMAY 102

Db 118 DQSSSLIHNPTGTNF-LELIGHISIDFIHCDOELQDALTRPNLUSKKLAEAPERH 176

Qy 103 -----LSENVSKHIGLSQELIGHISIDFIHCDOELQDALTPQQTLSRRKVAPTER 156

Db 177 FPSLRMKSTLTSRGRTLNKAATWKVLUHCSCMHMRAYKPAOTS PAGSPRSEPPQCLVLIC 236

Qy 157 CFSLRMKSTLTSRGRTLNKAATWKVLUHCSCMHMRAYKPAOTS PAGSPDSEPLQCLVLIC 216

Db 237 CEATPOLPFHDGA 249

Qy 217 EAIP---HPGS 225

Db

RESULT 11

ID AAB94934 ID AAB94934 standard; protein; 648 AA.

AC AAB94934; XX DT 26-JUN-2001 (first entry)

DB Human protein sequence SEQ ID NO:16415.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX EP1074617-A2.

XX PD 07-FEB-2001.

XX PR 28-JUL-2000; 2000BP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

XX PR 27-AUG-1999; 99JP-00300253.

XX PR 11-JAN-2000; 2000JP-00118776.

XX PR 02-MAY-2000; 2000JP-00183777.

XX PR 09-JUN-2000; 2000JP-00241899.

XX PA (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T,

XX DR WPI: 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX

PS Claim 8; SEQ ID NO 16415; 2537PP + Sequence Listing; English.

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dt primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the 5602 oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH1366 to AAH1368 and AAH18742 represent human amino acid sequences; AAB9593 to AAB9599 represent human cDNA sequences; and AAH1362 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX Sequence 648 AA;

Query Match 35.0%; Score 573; DB 4; Length 648;  
Best Local Similarity 6.0%; Pred. No. 7.7e-50; Matches 125; Conservative 11; Mismatches 27; Indels 42; Gaps 5;

QY 57 MRITISYLKMLRCLAGGKGGRATGRGLLPEGGGFRGHTRGRHGLVPGKQ---- 109  
Db 1 MRITISYLKMLRCLAGGKGGRATGRGLLPEGGGFRGHTRGRHGLVPGKQ---- 109  
QY 110 ---QAPGQSVDLCSHLIHNPNPGTNFS-LEHIGHSIFDFIPCDQBELQALTPRN 164  
Db 41 WVLTAEGMAY----ISENVSKGLGISOLENLIGHISIFDFIPCDQBELQALTPRN 94  
QY 165 LSKKKLEAPTERHSLRMKSTLSRGRTNLKAATWKVHCSHMRAYKPPAOTSPCSP 224  
Db 95 LSRRKVKEAPTERCFSLRMKSTLSRGRTNLKAATWKVHCSHMRAYKPAOTSPCSP 154

QY 225 RSEPLQLQCLVLCIAIPQLPFDGA 249  
Db 155 DSEPLQLQCLVLCIAIP---HPGS 175

RESULT 12  
ID AAL93710  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

AC AAL93710;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:13303.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

XX EP1074617-A2.  
XX 07-FEB-2001.  
XX 28-JUL-2000; 2000BP-00116126.

XX

PA (HBLI-) HELIX RES INST.

XX PR 29-JUL-1999; 99JP-00248936.  
PR 27-AUG-1999; 99JP-00300253.  
PR 11-JAN-2000; 2000JP-00118767.  
PR 02-MAY-2000; 2000JP-00183767.  
PR 09-JUN-2000; 2000JP-00241899.

XX DR WPI; 2001-318749/34.

XX PS Claim 8; SEQ ID NO 13303; 2537PP + Sequence Listing; English.

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

CC Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

CC (a) an oligo-dt primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the 5602 oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primers can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH1366 to AAH1368 and AAH18742 represent human amino acid sequences; AAB9593 to AAB9599 represent human cDNA sequences; and AAH1362 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX Sequence 199 AA;

Query Match 31.7%; Score 518.5; DB 4; Length 199;  
Best Local Similarity 6.0%; Pred. No. 6.2e-45; Matches 104; Conservative 9; Mismatches 13; Indels 27; Gaps 2;

QY 120 CTSSSLIRNPPTGTFNPF----SLELIGHISIFDFIPCDQBELQ 156  
Db 14 CTSWLTRCPSPAAASAPTWTRPLSCASPSATCACTASAPQLELIGHISIFDFIPCDQBELQ 73

QY 157 DALTPRENTLKKEAPTERHSLRMKSTLSRGRTNLKAATWKVHCSHMRAYKPPA 216  
Db 74 DALTPQQTLSRKKVEAPTERCFSLRMKSTLSRGRTNLKAATWKVHCSHMRAYKPA 133

QY 217 QTPSPAGSRPSEPLQLQCLVLCIAIPQLPFDGA 249  
Db 134 QTPSPAGSRDSEPLQLQCLVLCIAIP---HPGS 162

RESULT 13  
ID AAL6229  
XX ADI6229 standard; protein; 513 AA.  
AC AAL6229;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Human nucleic acid-associated protein (NAAP) #14.  
XX  
KW human; nucleic acid-associated protein; NAAP; autoimmune disorder;



quantifying differential gene expression, can be used in gene therapy, to formulate prognosis and to design a treatment regimen and to monitor the efficacy of treatment. The present sequence represents the amino acid sequence of a protein encoded by a cDNA differentially expressed in a liver disorder.